Lecture Notes in Networks and Systems 910

Radek Silhavy
Petr Silhavy *Editors*

Data Analytics in System Engineering

Proceedings of 7th Computational Methods in Systems and Software 2023, Vol. 3

Springer

Lecture Notes in Networks and Systems 910

Series Editor

Janusz Kacprzyk, *Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland*

Advisory Editors

Fernando Gomide, *Department of Computer Engineering and Automation—DCA, School of Electrical and Computer Engineering—FEEC, University of Campinas—UNICAMP, São Paulo, Brazil*
Okyay Kaynak, *Department of Electrical and Electronic Engineering, Bogazici University, Istanbul, Türkiye*
Derong Liu, *Department of Electrical and Computer Engineering, University of Illinois at Chicago, Chicago, USA*
Institute of Automation, Chinese Academy of Sciences, Beijing, China
Witold Pedrycz, *Department of Electrical and Computer Engineering, University of Alberta, Alberta, Canada*
Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland
Marios M. Polycarpou, *Department of Electrical and Computer Engineering, KIOS Research Center for Intelligent Systems and Networks, University of Cyprus, Nicosia, Cyprus*
Imre J. Rudas, *Óbuda University, Budapest, Hungary*
Jun Wang, *Department of Computer Science, City University of Hong Kong, Kowloon, Hong Kong*

The series "Lecture Notes in Networks and Systems" publishes the latest developments in Networks and Systems—quickly, informally and with high quality. Original research reported in proceedings and post-proceedings represents the core of LNNS.

Volumes published in LNNS embrace all aspects and subfields of, as well as new challenges in, Networks and Systems.

The series contains proceedings and edited volumes in systems and networks, spanning the areas of Cyber-Physical Systems, Autonomous Systems, Sensor Networks, Control Systems, Energy Systems, Automotive Systems, Biological Systems, Vehicular Networking and Connected Vehicles, Aerospace Systems, Automation, Manufacturing, Smart Grids, Nonlinear Systems, Power Systems, Robotics, Social Systems, Economic Systems and other. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution and exposure which enable both a wide and rapid dissemination of research output.

The series covers the theory, applications, and perspectives on the state of the art and future developments relevant to systems and networks, decision making, control, complex processes and related areas, as embedded in the fields of interdisciplinary and applied sciences, engineering, computer science, physics, economics, social, and life sciences, as well as the paradigms and methodologies behind them.

Indexed by SCOPUS, INSPEC, WTI Frankfurt eG, zbMATH, SCImago.

All books published in the series are submitted for consideration in Web of Science.

For proposals from Asia please contact Aninda Bose (aninda.bose@springer.com).

Radek Silhavy · Petr Silhavy
Editors

Data Analytics in System Engineering

Proceedings of 7th Computational Methods in Systems and Software 2023, Vol. 3

Springer

Editors
Radek Silhavy
Faculty of Applied Informatics
Tomas Bata University in Zlin
Zlin, Czech Republic

Petr Silhavy
Faculty of Applied Informatics
Tomas Bata University in Zlin
Zlin, Czech Republic

ISSN 2367-3370 ISSN 2367-3389 (electronic)
Lecture Notes in Networks and Systems
ISBN 978-3-031-53551-2 ISBN 978-3-031-53552-9 (eBook)
https://doi.org/10.1007/978-3-031-53552-9

© The Editor(s) (if applicable) and The Author(s), under exclusive license
to Springer Nature Switzerland AG 2024

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Paper in this product is recyclable.

Preface

Welcome to Volume 1 of the conference proceedings for the esteemed Computational Methods in Systems and Software 2023 (CoMeSySo). This volume, titled "Software Engineering Methods in Systems and Network Systems," encapsulates the innovative strides and groundbreaking research presented by experts, scholars, and professionals from around the globe.

In today's digital age, the role of software engineering in shaping the future of systems and network systems cannot be understated. The papers and articles contained within this volume delve deep into the methodologies, practices, and tools that are at the forefront of this dynamic field. From novel approaches to software development to the optimization of network systems, the breadth and depth of topics covered here are a testament to the vibrant and evolving nature of software engineering.

The CoMeSySo conference has always been a melting pot of ideas, fostering collaborations, and discussions that push the boundaries of what's possible in computational methods. This year, we were privileged to witness a confluence of minds, all dedicated to advancing the state of the art in software engineering for systems and network systems.

We want to extend our heartfelt gratitude to all the authors, reviewers, and members of the organizing committee. Their dedication, hard work, and passion have made this volume not just a collection of papers but a beacon for future research and development.

To our readers, we hope this volume serves as both an inspiration and a resource. Whether you are a seasoned professional, a budding researcher, or a curious enthusiast, the insights and knowledge shared within these pages will enrich your understanding and fuel your passion for software engineering.

Thank you for being a part of this journey. We look forward to the continued growth and evolution of the CoMeSySo community and to the innovations that the future holds.

Radek Silhavy
Petr Silhavy

Organization

Program Committee

Program Committee Chairs

Petr Silhavy	Tomas Bata University in Zlin, Faculty of Applied Informatics
Radek Silhavy	Tomas Bata University in Zlin, Faculty of Applied Informatics
Zdenka Prokopova	Tomas Bata University in Zlin, Faculty of Applied Informatics
Roman Senkerik	Tomas Bata University in Zlin, Faculty of Applied Informatics
Roman Prokop	Tomas Bata University in Zlin, Faculty of Applied Informatics
Viacheslav Zelentsov	Doctor of Engineering Sciences, Chief Researcher of St. Petersburg Institute for Informatics and Automation of Russian Academy of Sciences (SPIIRAS)
Roman Tsarev	Department of Information Technology, International Academy of Science and Technologies, Moscow, Russia
Stefano Cirillo	Department of Computer Science, University of Salerno, Fisciano (SA), Italy

Program Committee Members

Juraj Dudak	Faculty of Materials Science and Technology in Trnava, Slovak University of Technology, Bratislava, Slovak Republic
Gabriel Gaspar	Research Centre, University of Zilina, Zilina, Slovak Republic
Boguslaw Cyganek	Department of Computer Science, University of Science and Technology, Krakow, Poland
Krzysztof Okarma	Faculty of Electrical Engineering, West Pomeranian University of Technology, Szczecin, Poland

Monika Bakosova	Institute of Information Engineering, Automation and Mathematics, Slovak University of Technology, Bratislava, Slovak Republic
Pavel Vaclavek	Faculty of Electrical Engineering and Communication, Brno University of Technology, Brno, Czech Republic
Miroslaw Ochodek	Faculty of Computing, Poznan University of Technology, Poznan, Poland
Olga Brovkina	Global Change Research Centre Academy of Science of the Czech Republic, Brno, Czech Republic & Mendel University of Brno, Czech Republic
Elarbi Badidi	College of Information Technology, United Arab Emirates University, Al Ain, United Arab Emirates
Luis Alberto Morales Rosales	Head of the Master Program in Computer Science, Superior Technological Institute of Misantla, Mexico
Mariana Lobato Baes	Research-Professor, Superior Technological of Libres, Mexico
Abdessattar Chaâri	Laboratory of Sciences and Techniques of Automatic Control & Computer Engineering, University of Sfax, Tunisian Republic
Gopal Sakarkar	Shri. Ramdeobaba College of Engineering and Management, Republic of India
V. V. Krishna Maddinala	GD Rungta College of Engineering & Technology, Republic of India
Anand N. Khobragade (Scientist)	Maharashtra Remote Sensing Applications Centre, Republic of India
Abdallah Handoura	Computer and Communication Laboratory, Telecom Bretagne, France
Almaz Mobil Mehdiyeva	Department of Electronics and Automation, Azerbaijan State Oil and Industry University, Azerbaijan

Technical Program Committee Members

Ivo Bukovsky, Czech Republic
Maciej Majewski, Poland
Miroslaw Ochodek, Poland
Bronislav Chramcov, Czech Republic
Eric Afful Dazie, Ghana
Michal Bliznak, Czech Republic

Donald Davendra, Czech Republic
Radim Farana, Czech Republic
Martin Kotyrba, Czech Republic
Erik Kral, Czech Republic
David Malanik, Czech Republic
Michal Pluhacek, Czech Republic
Zdenka Prokopova, Czech Republic
Martin Sysel, Czech Republic
Roman Senkerik, Czech Republic
Petr Silhavy, Czech Republic
Radek Silhavy, Czech Republic
Jiri Vojtesek, Czech Republic
Eva Volna, Czech Republic
Janez Brest, Slovenia
Ales Zamuda, Slovenia
Roman Prokop, Czech Republic
Boguslaw Cyganek, Poland
Krzysztof Okarma, Poland
Monika Bakosova, Slovak Republic
Pavel Vaclavek, Czech Republic
Olga Brovkina, Czech Republic
Elarbi Badidi, United Arab Emirates

Organizing Committee Chair

Radek Silhavy — Tomas Bata University in Zlin, Faculty of Applied Informatics
email: comesyso@openpublish.eu

Conference Organizer (Production)

Silhavy s.r.o.
Web: https://comesyso.openpublish.eu
Email: comesyso@openpublish.eu

Conference Website, Call for Papers

https://comesyso.openpublish.eu

Contents

Interpretable Rules with a Simplified Data Representation - a Case Study with the EMBER Dataset 1
Ján Mojžiš and Martin Kenyeres

Spatial Econometric Models: The Pursuit of an Accurate Spatial Structure with an Application to Labor Market 11
Tomáš Formánek

Use of Neighborhood-Based Bridge Node Centrality Tuple for Preferential Vaccination of Nodes to Reduce the Number of Infected Nodes in a Complex Real-World Network 27
Natarajan Meghanathan, Kapri Burden, and Miah Robinson

Social Media Applications' Privacy Policies for Facilitating Digital Living 37
Kagiso Mphasane, Vusumuzi Malele, and Temitope Mapayi

Correlation Analysis of Student's Competencies and Employ-Ability Tracer Study of Telkom University Graduates 49
P. H. Gunawan, I. Palupi, Indwiarti, A. A. Rohmawati, and A. T. Hanuranto

Hardware Design and Implementation of a Low-Cost IoT-Based Fire Detection System Prototype Using Fuzzy Application Methods 61
Emmanuel Lule, Chomora Mikeka, Alexander Ngenzi, Didacienne Mukanyiligira, and Parworth Musdalifah

Closing the Data Gap: A Comparative Study of Missing Value Imputation Algorithms in Time Series Datasets 77
Sepideh Hassankhani Dolatabadi, Ivana Budinská, Rafe Behmaneshpour, and Emil Gatial

Motivation to Learn in an E-learning Environment with Fading Mark 91
Roman Tsarev, Younes El Amrani, Shadia Hamoud Alshahrani, Naim Mahmoud Al Momani, Joel Ascencio, Aleksey Losev, and Kirill Zhigalov

Hub Operation Pricing in the Intermodal Transportation Network 100
Alexander Krylatov and Anastasiya Raevskaya

A New Approach to Eliminating the "Flip" Effect of the Approximating Function Under Conditions of a Priori Uncertainty 112
V. I. Marchuk, A. A. Samohleb, M. A. Laouar, and H. T. A. Al-Ali

Modeling the Alienability of an Electronic Document 119
Alexander V. Solovyev

Automatic Generation of an Algebraic Expression for a Boolean Function in the Basis $\{\wedge, \vee, \neg\}$ 128
Roman Tsarev, Roman Kuzmich, Tatyana Anisimova, Biswaranjan Senapati, Oleg Ikonnikov, Viacheslav Shestakov, Alexander Pupkov, and Svetlana Kapustina

Applying a Recurrent Neural Network to Implement a Self-organizing Electronic Educational Course 137
Ruslan Khakimzyanov, Sadaquat Ali, Bekbosin Kalmuratov, Phuong Nguyen Hoang, Andrey Karnaukhov, and Roman Tsarev

Empowering Islamic-Based Digital Competence and Skills: How to Drive It into Reconstructing Safety Strategy from Gender Violence 146
Miftachul Huda, Mukhamad Hadi Musolin, Anassuzastri Ahmad, Andi Muhammad Yauri, Abu Bakar, Muhammad Zuhri, Mujahidin, and Uswatun Hasanah

From Digital Ethics to Digital Community: An Islamic Principle on Strengthening Safety Strategy on Information 165
Miftachul Huda, Mukhamad Hadi Musolin, Mohamad Hazli Ismail, Andi Muhammad Yauri, Abu Bakar, Muhammad Zuhri, Mujahidin, and Uswatun Hasanah

Cyber Security Management in Metaverse: A Review and Analysis 183
Farnaz Farid, Abubakar Bello, Nusrat Jahan, and Razia Sultana

Enhancing Educational Assessment: Leveraging Item Response Theory's Rasch Model 194
Georgi Krastev, Valentina Voinohovska, and Vanya Dineva

Particular Analysis of Regression Effect Sizes Applied on Big Data Set 203
Tomas Barot, Marek Vaclavik, and Alena Seberova

Applied Analysis of Differences by Cross-Correlation Functions 210
Tomas Barot, Ladislav Rudolf, and Marek Kubalcik

Paraphrasing in the System of Automatic Solution of Planimetric Problems 217
Sergey S. Kurbatov

The Formation of Ethno-Cultural Competence in Students Through the Use of Electronic Educational Resources 226
Stepanida Dmitrieva, Tuara Evdokarova, Saiyna Ivanova, Ekaterina Shestakova, Marianna Tolkacheva, and Maria Andreeva

A Model of Analysis of Daily ECG Monitoring Data for Detection of Arrhythmias of the "Bigeminy" Type 239
D. V. Lakomov, Vladimir V. Alekseev, O. H. Al Hamami, and O. V. Fomina

Development of a Data Mart for Business Analysis of the University's Economic Performance 245
Sergei N. Karabtsev, Roman M. Kotov, Ivan P. Davzit, Evgeny S. Gurov, and Andrey L. Chebotarev

Explainability Improvement Through Commonsense Knowledge Reasoning 259
HyunJoo Kim and Inwhee Joe

Approximability of Semigroups with Finiteness Conditions 278
Svetlana Korabelshchikova, Larisa Zyablitseva, Boris Melnikov, and Dang Van Vinh

Semigroup Invariants of Graphs with Respect to Their Approximability 286
Svetlana Korabelshchikova, Larisa Zyablitseva, Boris Melnikov, and Dang Van Vinh

Algorithmizing Aspects of Some Combinatorial Block-Designs Implemented in Network Systems 293
Alexander Frolov, Natalya Kochetova, and Anton Klyagin

Spectral Method of Analysis and Complex Processing of Video Information from Multi-Zone Images 330
Valery V. Vasilevskiy

Cyber-Physical Control System for Personal Protective Equipment Against Infectious Diseases Transmitted by Airborne Droplets 336
Mikhail Golosovskiy, Alexey Bogomolov, Eugene Larkin, and Tatiana Akimenko

Concept of Smart Personal Protection Equipment Against Infectious Diseases 343
Alexey Bogomolov, Eugene Larkin, and Tatiana Akimenko

Software for Urban Space Researching with Public Data for Irkutsk 352
O. A. Nikolaychuk, D. E. Kosogorov, and Yu. V. Pestova

Core-Intermediate-Peripheral Index: Factor Analysis of Neighborhood and Shortest Paths-Based Centrality Metrics 363
Natarajan Meghanathan

Multipole Engineering of Optical Forces 373
Denis Kislov and Vjaceslavs Bobrovs

Data Mart in Business Intelligence with Ralph Kimball for Commercial Sales 380
Alessandro Chanco Torres, Angel Quiñonez Gastelu, Juan J. Soria, Mercedes Vega Manrique, and Lidia Segura Peña

Intelligent Endoscopic Examination of Internal Openings for Drilling Quality Control 397
Anton Ivaschenko, Vladimir Avsievich, Vera Turkova, Andrey Belikov, and Natalia Chertykovtseva

Foreign Languages Teaching Technologies and Methods in Traditional and Distance Forms of Learning 404
E. D. Lavrinenko, L. A. Bayrak, J. I. Erkenova, I. Sh. Kappusheva, O. A. Frolova, and N. A. Bystrov

The Usage of Distance Educational Technologies for Learning English 411
V. A. Sogrina, D. A. Stanovova, J. I. Erkenova, I. Sh. Kappusheva, F. A. Nanay, and A. V. Filipskaya

Evaluating the Effectiveness of Flipped Classrooms Using Linear Regression 418
Roman Tsarev, Biswaranjan Senapati, Shadia Hamoud Alshahrani, Alsu Mirzagitova, Shokhida Irgasheva, and Joel Ascencio

Building a Digital Twin of a Social Media User Based on Implicit Profile Data 428
Ilya Shirokov, Zulfiya Kamaldinova, Irina Dubinina, and Anton Ivaschenko

Analysis of a Data Set to Determine the Dependence of Airline Passenger Satisfaction 434
V. S. Tynchenko, Borodulin, I. I. Kleshko, V. A. Nelyub, and Rukosueva

Author Index 459

Interpretable Rules with a Simplified Data Representation - a Case Study with the EMBER Dataset

Ján Mojžiš(✉) and Martin Kenyeres

Institute of Informatics, Slovak Academy of Sciences, Dúbravská Cesta 9, Bratislava, Slovakia
{jan.mojzis,martin.kenyeres}@savba.sk

Abstract. The EMBER dataset is a popular malware samples dataset open to researchers to evaluate. Despite its wide coverage in various proposals and papers, there are no evaluations of EMBER with a simplified (binary) data representation. In this paper, we propose 1) a simplified data representation of EMBER and 2) a proprietary feature selection method called ESFS (Ember Special Feature Selection) tailored to this representation. We demonstrate that simplified data representation is a viable alternative to the original representation and that the ESFS method improves the performance of all evaluated models above the conventional methods (Pearson correlation and information gain). Lastly, we use the best-performing model (J48) to extract and visualise several example rules for malware classification.

Keywords: malware classification · rule-based models · tree visualisation · feature selection · binary representation · EMBER dataset

1 Introduction

Malware detection is traditionally (in antiviral software) based on signature matching where the testing sample is matched to the signatures in the malware signatures dataset. While this approach can be quick and efficient, it fails to detect harmful malware samples, provided that signatures are not present in the dataset. Another approach to counter this disadvantage is to use dynamic analysis, where an evaluation is performed towards potentially harmful content during its execution time, in order to identify whether it does something suspicious or harmful. Despite the absence of signatures, static analysis can still be used, but some form of generalisation is required to compensate for signatures absence. For the scope of this paper, we exclude dynamic analysis and discuss only static analysis approach.

Employing a generalisation in the static analysis process, detection could be extended to match previously overlooked harmful samples. Our motivation here is our recent work on the ontology creating for use in concept learning with use of description logics learning algorithms, where concepts are created and are used instead of signatures matching [1]. A generalisation in this context means not to use direct features gathered from the malware

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 1–10, 2024.
https://doi.org/10.1007/978-3-031-53552-9_1

dataset, but to use refined (newly calculated) features based on the original features. For this purpose, in this paper, we evaluate the EMBER dataset with several traditional machine learning models (namely, J48, Decision table, and JRip (a RIPPER algorithm implementation for WEKA)) with the WEKA software. From the original feature set of vector features, we refined a set of more than 9,000 features. In our former work, we introduced the rules generation based on information gain feature selection and the J48 tree model [2]. In this paper, we perform the feature selection with the following methods: Pearson correlation (PC), information gain (IG) and our own proprietary method Ember Special Feature Selection (ESFS), which was created particularly for the use with the EMBER dataset. Our results indicate that ESFS outperforms traditionally used methods of feature selection with all evaluated models.

Based on the results (mainly lowest false positive and highest true positive rates for malware samples), we extract interpretable rules and generate tree visualisation with several shortest paths from the J48 model. The rules from the resulting J48 model could be used in the concept-learning process.

The rest of this paper is organised as follows: the next section is dedicated to related works, followed by Methodology and Results and discussion. The paper concludes with Conclusion.

2 Related Works

Various evaluations and proposals have been conducted since the EMBER dataset was published. In this paper, we focus on the interpretability of malware classification.

Dolejš and Jureček focus on the interpretability of machine learning models' results using decision lists generated by two rule-based classifiers, I-REP, and RIPPER [8] with the EMBER dataset. Contrary to our approach (with binary representation), they use original values of the attributes (numeric). For feature space reduction, they use PCA and Random Forest with the resulting feature count 200.

Trizna [3] performs a dynamic analysis with the malware-oriented Windows kernel emulator and achieves higher analysis rates compared to virtualization. He proposes a composite solution with multiple pre-trained modules and meta-model instead of a single feature vector. His acquired dataset consists of more than 100k executable samples, however, he intentionally omitted dynamic link library (DLL) PE files. His hybrid analysis yields improved performance, especially under low false-positive requirements. However, his models are not interpretable (he uses a combination of neural networks (FFNN and CNN)).

Aggarwal et al. [4] perform feature extraction/selection on the EMBER dataset. The result is rather radical feature space reduction however, the original two-class samples are split into several categories (banking, trojan, adware, spyware). The evaluations are performed with the use of CNN and their proposed selective targeted transfer learning (STTL). The resulting accuracy is above 90%, and the byte entropy category is ranked at the top with 9 ranking features in Trojan, 5 with Adware, and 11 features in Spyware family.

Oyama et al. [5] select EMBER features based on the LightGBM algorithm. The authors evaluate accuracy, learning time, and data size. From the EMBER feature groups,

by overall score, they select as the best general, header, and strings. In order to reach at least 90% of accuracy, no single feature group is enough despite a single feature group containing more than a hundred features; therefore, several feature groups are required. The features and the feature groups they select are based on the LightGBM algorithm results. They also consider the possibility to use non-machine learning methods, like chi square test (in this paper, we use information gain and Pearson correlation).

Kumar and Geetha [6] focus on a low-end classifier design XGboost with the objective of the removal of the noisy features and hyperparameter tuning. Using 800K EMBER dataset samples, they achieved nearly 100% accuracy (98.2%), but the feature space includes 2,351 features. Although the model does not overfit (due to various regularisation steps), the results are not interpretable.

3 Methodology

3.1 Experiments

All of the experiments, except for JRip, were run on a single computer platform (Intel I7 @ 2.80 GHz) with 16 GB of RAM running MS Windows 7. JRip was evaluated on a single computer platform (I9-10980XE @ 3.00 GHz) with 256 GB RAM.

3.2 Dataset Description and Preprocessing

Our experiments are conducted on the publicly available dataset EMBER (Elastic Malware Benchmark for Empowering Researchers) [7]. We used the most current version released in 2018. With the use of static analysis, authors extracted features from binaries of monetized software and incorporated them into the dataset. The EMBER dataset is divided into a training set with 900K samples (300K malware, 300K benign, and 300K unlabeled) and a test set with 200K samples (100K malware and 100K benign). We ignore unlabeled samples in our experiments and use the train set of 300K malware and 300k benign samples, which we further divided into 80% train and 20% test subsets.

The format of the EMBER is JSON. For each sample, eight groups of raw features are present. However, contrary to Dolejš [8], we do not use the original features present in the dataset. Instead, we calculate derived features from the original ones. We binarize features (each feature can be either true or false - boolean, denoting the presence or the absence of a feature) to create a simplified representation. This representation covers our general objective to ontology creation and the use of conceptual learning in the process of malware detection [1]. Actually, our representation of the features is a variation of Švec et al. [9].

Several features are calculated, e.g., *has_high_entropy*. This feature relates to the section entropies. If there (in a sample) is any such section where entropy is high (7 and more), *has_high_entropy* is set to true. Together, we collect features from various feature groups (sections, general, imports, actions, about the presence of some property). In the result, we collected 9.606 distinct features. We focus more on the functional characteristics of the code samples (we put no importance on strings or histogram parts, but we incorporate the information on actions instead – e.g. “kill window” and imports). We use three methods in order to reduce feature space.

3.3 Tree Visualisation

To visualise trees for this paper, we use https://graphviz.org/ for Java. To extract rules, we parse the textual representation of the J48 tree model generated by WEKA and reconstruct the data structure of branches and leaves. During this process, we put several conditions a) short paths are chosen (fewer rules), b) leaves cover most samples of the dataset, and 3) the best accuracy is required (high values of true positives and lower values of false positives).

3.4 Feature Selections Methods

In order to reduce feature space, prior to the dataset evaluation, we use three different feature selection methods: IG, PC, and the third is our own proprietary method ESFS (Eq. 1). We evaluate 25, 50, 75, 100, and 200 of the best-ranked features based on a particular selection method.

$$\sum_{i=1}^{m} \frac{\left|\sum_{i}^{n}(a_i == 0) - \sum_{1}^{n}(a_i == 1)\right|}{n} \tag{1}$$

where

$a_i \in A = \{a_1, a_2, a_3, \ldots, a_m\}$.

m - attributes count.

n - records count.

then descending sort for all a_i values.

The principle of ESFS is to enhance the contrast among the attributes. Thus, we seek attributes that are important for the first class but not for the second one (and vice versa).

Overall, there are two classes. The mentioned contrast can be understood as follows: attributes that are in 50% of cases of the first class and in 50% of the cases of the second class are considered the least important. In contrast, those that are in 90% of the first class and in 10% of the second class (and vice versa) are found significant.

3.5 Evaluation Setup

Dataset evaluation is performed with the WEKA 3.8.4 platform (an open-source Machine learning environment) with simple models J48 (implementation of C4.5 tree-based model), JRip (RIPPER algorithm proposed by Dolejš and Jureček [8]), and Decision table.We do not use boosting as it would counter the interpretability of the model. We evaluate the performance of models using true positive rate (TPR, Eq. 2), false positive rate (FPR, Eq. 3), and accuracy (ACC, Eq. 4), which are also used in [8].

$$FPR = \frac{FP}{FP + TN} \tag{2}$$

$$TPR = \frac{TP}{TP + FN} \tag{3}$$

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

where

TP - true positive, correctly predicted malicious samples as malicious
TN - true negative, correctly predicted benign samples as benign
FP - false positive, incorrectly predicted benign samples as malicious
FN - false negative, incorrectly predicted malicious samples as benign.

4 Results and Discussion

4.1 Evaluation of Models

Three different models were evaluated with three feature selection methods and five different feature counts. In the first iteration, we focused on the lowest possible FPR. Table 1 lists evaluated models with their parameters and Fig. 1 compares these models by their resulting FPR. The feature count is 25. As seen from the results, the worst model here is JRip, and the best is J48 with our ESFS feature selection method. ESFS feature selection, here, gives the lowest FPR for all models.

Table 1. Evaluated models.

Model	Short name	Description
Decision table	DT	WEKA default
Decision table	DT_RMSE	Evaluates with RMSE
Decision table	DT_iBk	Use iBk
Decision table	DT_iBk_RMSE	Use iBk and evaluates with RMSE
J48	J48	WEKA default
JRip	JRip	RIPPER algorithm, WEKA default

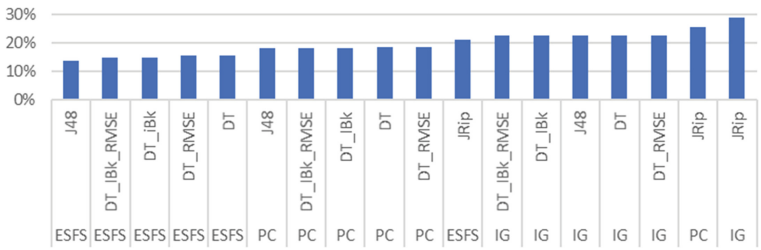


Fig. 1. FPR, lowest is the best. Three different models, three feature selection methods and 25 features. Decision table had several parameters modified according to Table 1 (Short name denotation).

Additionally, we evaluated models for 50, 75, 100 and 200 feature counts. Due to the performance issues with JRip model with 100 and more features, we decided to exclude it from the further evaluations. Based lower FPR values of DT in Fig. 1, for further evaluations, we use the configuration DT_iBk_RMSE. Figure 2 compares the

remaining two models for 50 features and above. The best model is still J48 with our ESFS feature selection method. FPR for decision table and J48 models is measured in Fig. 3.

Regarding the feature count, we decided to use as many as 200 features, which is the same feature count as Dolejš and Jurečen proposed [8]. With J48, ESFS, and 200 features, we gathered the following results for the malware class: accuracy: 91.61%, TPR: 92.3%, and FPR: 9.1%.

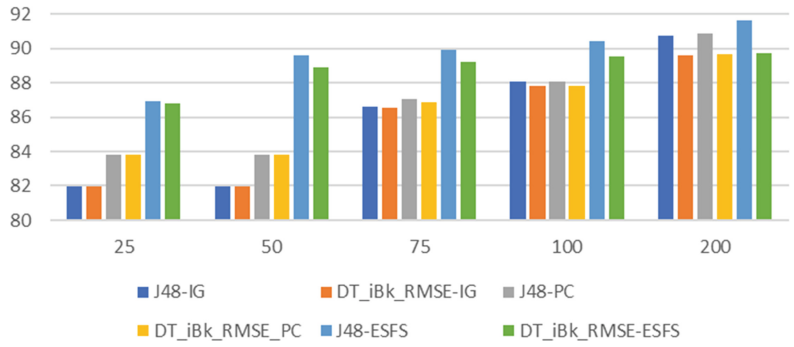


Fig. 2. Accuracy, highest is the best. Two models, three feature selection methods and five feature counts.

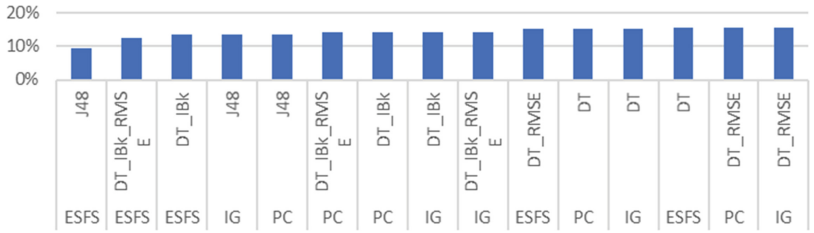


Fig. 3. FPR, lowest is the best. Two different models, three feature selection methods and 100 features.

4.2 Extracted Rules

The resulting J48 tree model supplied with 200 features during the train phase is using only 124 of them. The most important feature (according to our ESFS method) in all paths generated by this J48 tree model (80% of the original EMBER train set) is the *is_dll* feature and is contained in all the paths. This feature controls whether a code sample is a dynamic link library (DLL) or not. If it has a value of 0, it is not a DLL. In the dataset, there are 45.8% benign DLL samples and only 6.3% malware DLL samples. The second most important feature is the *has_section_high_entropy*. This feature controls whether the sample (malware or benign) has any section where entropy is high. Actually, this is our generated feature and is based on section entropies of the original EMBER

sample. If there is any section in the sample where entropy is high (value $> = 7$), the *has_section_high_entropy* feature is set to 1 for a given sample. In the dataset, only 17.1% benign samples have high entropy, while up to 51.2% of malware samples have their entropy high.

Table 2 lists the rules based on the shortest paths for benign and malware samples. The shortest path for benign covers 1,745 samples, of which 12 are false positive (e.g., they are malware actually), and the path length is 4 (four features are used). The shortest Malware path covers 7,611 samples, all of which are covering malware samples, and the path length is 4.

The highest coverage path for benign samples covers 63,441 samples out of which 459 are false positives (malware), and the path length is 15. The highest coverage path for malware samples covers 14,170 samples out of 62 are false positive (benign), and the path length is 27.

We can use tree visualisation (WEKA's GraphViz plugin, controlled with API) to visualise several shorter paths which have low FPR (Fig. 4).

Table 2. Extracted rules with the priority on shortest paths. Two classes: benign and malware.

Benign	Malware
is_dll = 0	is_dll = 0
sect_pdata_has_CNT_INITIALIZED_DATA = 0	sect_pdata_has_CNT_INITIALIZED_DATA = 0
sect_didat_writable = 1	sect_didat_writable = 0
has_write_execute_section = 0	sect_coderpub_readable = 1

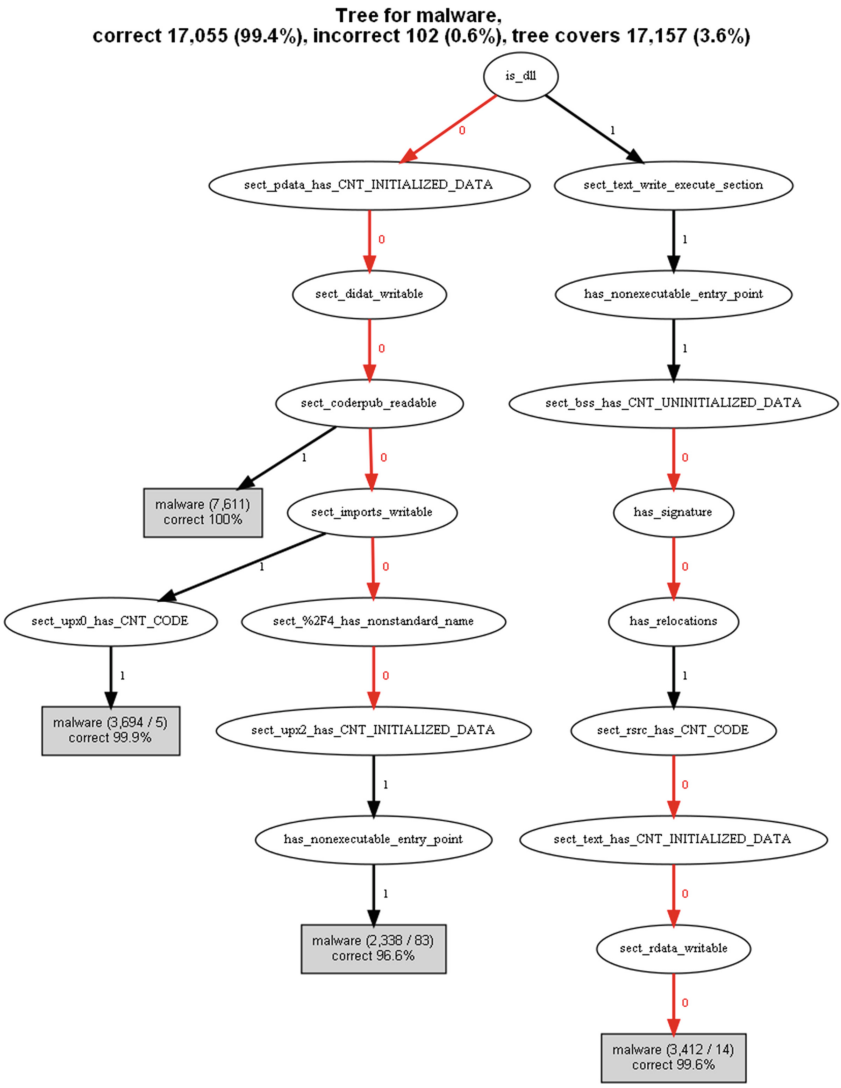


Fig. 4. Several shortest paths for malware samples, covering in total 17,157 samples (3.6%), out of which 102 (0.6%) are false positives.

5 Conclusion

In this paper, we propose an approach to use rule and tree-based models to extract interpretable rules for malware classification. We evaluated several models (J48, JRip, and Decision table) supplied with the features selected by three different feature selection methods. Dataset is the EMBER dataset. As the best model, we select J48 tree-based model with our novel ESFS feature selection method according to its performance (the

lowest false positives and highest accuracy). With the use of J48, we extracted (Table 2) and visualised (Fig. 4) several example malware classification rules.

Despite the high performance of many case study models dealing with the EMBER dataset [5, 6], their results are not interpretable, which is our vital objective. Even if interpretability is possible [4], in order to incorporate results for conceptual learning, no rules could be generated in this case. For interpretable results of [8], not every case of high accuracy shows the lowest possible FPR (Table 2 and 4 in [8]). In addition, we must consider our goal to use the rules in conceptual learning. We need a simplified binary representation of the EMBER data. Thus, even less performance of the model is compensated by its interpretability and simplicity tailored for conceptual learning.

We developed and evaluated a novel feature selection method Ember Special Feature Selection (ESFS), which outperforms other two traditionally used methods: information gain (IG) and Pearson correlations (PC). This method was also evaluated in the dataset from another domain (sentiment analysis), but the performance significance was not confirmed. It would be suitable to perform supplemental evaluations of this method within the malware classification domain.

In future work, we prepare to incorporate generated rules from the J48 tree model with the conceptual learning with the objective of improvement of the malware classification using conceptual learning and descriptive logics.

Acknowledgments. This work was supported by the following grants VEGA no. 2/0135/23 "Intelligent sensor systems and data processing" and APVV no. APVV-19-0220.

References

1. Svec, P., Balogh, S., Homola, M.: Experimental evaluation of description logic concept learning algorithms for static malware detection. In: ICISSP 2021, pp. 792–799 (2021)
2. Mojžiš, J.: On the possibility of interpretable rules generation for the classification of malware samples. Industry 4.0. **7**(6), 248–250 (2022)
3. Trizna, D.: Quo Vadis: hybrid machine learning meta-model based on contextual and behavioral malware representations. In: Proceedings of the 15th ACM Workshop on Artificial Intelligence and Security, pp. 127–136, 11 November 2022
4. Aggarwal, P., Ahamed, S.F., Shetty, S., Freeman, L.J.: Selective targeted transfer learning for malware classification. In: 2021 Third IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications (TPS-ISA), pp. 114–120. IEEE, 13 December 2021
5. Oyama, Y., Miyashita, T., Kokubo, H.: Identifying useful features for malware detection in the ember dataset. In: 2019 seventh international symposium on computing and networking workshops (CANDARW), pp. 360–366. IEEE, 26 November 2019
6. Kumar, R., Geetha, S.: Malware classification using XGboost-Gradient boosted decision tree. Adv. Sci. Technol. Eng. Syst. **5**, 536–549 (2020)
7. Hyrum, S.: Anderson and Phil Roth. EMBER: an open dataset for training static PE malware machine learning models. ArXiv e-prints, 2018

8. Dolejš, J., Jureček, M.: Interpretability of machine learning-based results of malware detection using a set of rules. In: Stamp, M., Aaron Visaggio, C., Mercaldo, F., Di Troia, F. (eds.) Artificial Intelligence for Cybersecurity. Advances in Information Security, vol. 54, pp. 107–136. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-97087-1_5
9. Švec, P., Balogh, Š., Homola, M., Kľuka, J.: Knowledge-based dataset for training PE malware detection models (2022). arXiv preprint arXiv:2301.00153

Spatial Econometric Models: The Pursuit of an Accurate Spatial Structure with an Application to Labor Market

Tomáš Formánek(✉)

Prague University of Economics and Business, Prague, Czech Republic
formanek@vse.cz
http://www.vse.cz

Abstract. In most applications of spatial econometrics, spatial prior information is used to distinguish close units (interacting, spatially dependent) from distant units (mutually independent). However, the true spatial setup is not actually known in most cases. Fortunately, even imperfect yet mostly valid spatial structures may lead to reasonably accurate model estimates. Still, the general validity of any used neighborhood structure should be consistently assessed as deviations from true setups cannot be easily avoided. Such assessments are typically based on statistical measures, which in turn facilitate and motivate the quest of improving spatial structures used for estimation. Enhanced spatial information may lead to higher efficiency of marginal effect estimation and better prediction accuracy. This article presents a heuristic algorithm that uses sectoral macroeconomic to generate enhanced neighborhood definitions. Suitable methods for statistical inference and efficiency verification are also provided. The proposed method is tailored for short panels and maximum likelihood estimators, yet its principles are generally applicable. An empirical demonstration of the approach is provided, based on NUTS2-level data for 10 contiguous EU member states.

Keywords: spatial structure · regional interactions · spillover effects · labor market

1 Introduction

State of the art regional macroeconomic analyses are becoming increasingly reliant on spatial quantitative methods and spatial econometrics [11]. This article addresses important features of regional panel data models, where the underlying assumption of cross-sectional independence is violated through spatial interactions [2,3].

In principle, spatial dependency resembles autoregressive processes present in time series of economic variables. In both cases, interactions are defined in terms of distance. However, there is one important distinction related to the 2D character of spatial processes. While the time axis is one-dimensional and we

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 11–26, 2024.
https://doi.org/10.1007/978-3-031-53552-9_2

only deal with unidirectional causal relationships, spatial processes are different: individual units are mutually dependent.

The functional form of spatial dependency is not directly observable in terms of the maximum distance and/or the distance-based decay in strength of spatial interactions [5]. Spatial dependency (its strength) can be described as a non-continuous function that follows from distance between units [13]. Often, spatial dependency functions are asymmetric (i.e. dependent on origin and destination) and prone to local structural breaks. Local discrepancies in spatial dependencies may even generate unintuitive effects in the global pattern of spatial dependency [17].

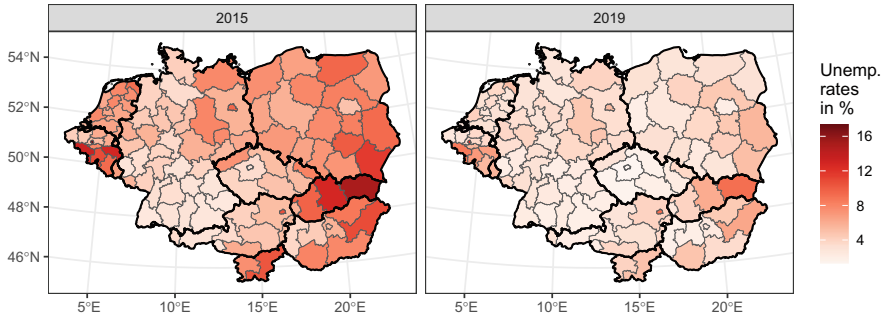


Fig. 1. Unemployment rates, years 2015 and 2019, NUTS2-level data shown for Austria, Belgium, Czechia, Denmark, Hungary, Luxembourg, the Netherlands, Poland, Slovenia and Slovakia

While spatial structures may be theoretically complex, empirical evidence often favours clear and prominent spatial dependency patterns that allow for identification of spatial patterns through statistical methods. As a simple illustration, Fig. 1 shows prominent and time-invariant spatial patterns in regional unemployment that can be visually identified even as observed data vary both across NUTS2 regions and across time.

Under a common assumption of spatial structures being time-invariant and pending on data availability of relevant sectoral geo-coded observations, statistical methods can be used to assess mutual dependencies and to evaluate potential discrepancies between distances and actual "accessibility" features. Many such methods can be used to assess, improve or even fully estimate spatial structures used in spatial econometric models [5,10].

This article presents a new algorithm that uses empirical sectoral economic data to enhance information based upon distance-based spatial structures. In practical applications, the algorithm proposed provides a substantial improvement in performance of spatial econometric models as compared to the traditional approach of generating underlying spatial structures based purely on spatial positions (distances, contiguity, etc.).

The rest of this contribution is structured as follows: next section describes spatial models and corresponding estimation and evaluation methods, section three covers the proposed method of improving spatial specification and thus model estimation. Section four features an illustrative application of the approach proposed, based on labor market data from selected EU countries. Section five concludes this contribution, along with the list of references.

2 Spatial Models: Definition, Estimation and Evaluation

Formulation and estimation of most cross sectional and panel-based spatial models starts by defining the underlying spatial structure. For economic applications, we typically deal with administrative units (countries or regions). Such spatial units are in fact polygons with a non-zero surface area and we can use their representative points – centroids – to measure distances between regions and to distinguish neighbors (close units) from distant and unrelated regions. Beside distances, the neighborhood structure can be established using a common border (contiguity) rule and through other specialized approaches [16].

For distance-base neighborhoods, identification of close and distant spatial units is often based on a connectivity matrix $\boldsymbol{C}$ such that

$$\boldsymbol{C} = [c_{ij}] = \begin{cases} 0 & \text{if} \quad i = j\,, \\ 0 & \text{if} \quad d_{ij} > \tau\,, \\ 1 & \text{if} \quad d_{ij} \leq \tau \quad \text{and} \quad i \neq j\,, \end{cases} \tag{1}$$

where $d_{ij} = d_{ji}$ is the aerial distance between two (geo-coded) spatial units i and j. For $c_{ij} = 1$, units i and j are neighbors, i.e. are sufficiently close to interact and influence each other mutually, and vice versa. Zeros on the main diagonal indicate that spatial units are not neighbors to themselves by definition. Parameter τ is a heuristically selected maximum distance threshold between two neighbors. Changes in τ can have a significant impact on the resulting $\boldsymbol{C}$-matrix, which is used as prior information for most estimating methods [12].

Symmetry of the $(N \times N)$ matrix $\boldsymbol{C}$, (i.e. $c_{ij} = c_{ji}$) is implied from the use of aerial distances d_{ij}. However, discrepancies between "accessibility" features and aerial distances can arise in various empirical applications [17]. This may lead to asymmetric connectivity matrices [4].

For contiguity-based patterns based on common borders, we set $c_{ij} = 1$ if regions i and j share a common border and vice versa. With the k-nearest neighbors (kNN) approach, each ith row of $\boldsymbol{C}$ contains k nonzero elements, selected by ranking distances from unit i: from (first) closest to kth closest. Such $\boldsymbol{C}$ matrices are generally non symmetric. The $\boldsymbol{C}$ matrix does not have to be binary: non-zero c_{ij} elements may reflect inverse distances.

Spatial econometric models are based on a transformation of the connectivity matrix $\boldsymbol{C}$ from (1): a spatial weights matrix $\boldsymbol{W}$ is calculated by row-wise standardization of $\boldsymbol{C}$. Each element of $\boldsymbol{W}$ is calculated as $w_{ij} = c_{ij} / \sum_{j=1}^{N} c_{ij}$ so that all row sums in $\boldsymbol{W}$ equal one: $\sum_{j=1}^{N} w_{ij} = 1$ for $\forall i$. Again, alternative standardization schemes can be applied for $\boldsymbol{W}$ construction [5].

2.1 Spatial Lag Model

An econometric spatial lag model is defined by spatial interactions in the dependent variable [4]. Nevertheless, the main discussion provided in this article may be easily extended to models encompassing spatial processes in their error terms and/or in the regressors [1,12]. For cross sectional data, the spatial lag model can be defined as

$$\boldsymbol{y} = \lambda \boldsymbol{W}\boldsymbol{y} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}, \tag{2}$$

where $\boldsymbol{y}$ is the $N \times 1$ dependent variable vector, $\boldsymbol{X}$ is the usual $N \times K$ matrix of regressors (includes the intercept element) with N being the number of observations and K the number of regressors (including intercept). Element $\boldsymbol{u}$ is the error term and $\boldsymbol{W}$ is the spatial weights matrix. Model parameter λ is a scalar describing the strength of spatial dependency and $\boldsymbol{\beta}$ is a vector of parameters (say, describing economic dynamics).

In most types of spatial models, the generalization from cross-sectional to panel data is relatively straight forward, very similar to the case of non-spatial models [14]. Model (8a)–(8b) is an example of such extension to panel data. Finally, it should be noted that $\boldsymbol{\beta}$-parameters are not marginal effects. However, using $\boldsymbol{W}$ and the estimates of λ and $\boldsymbol{\beta}$, marginal effects (direct and spillover) can be calculated along with their corresponding statistical inference [4].

2.2 Existing Estimation Approaches – Comparison of Methods

Under very general conditions [4], the maximum likelihood (ML) estimator may be used to produce estimates of all parameters of model (2): that is for $\boldsymbol{\beta}$, λ and random element variance σ^2. Assuming normal distribution of the error elements, ML function for Eq.(2) can be cast as

$$LL(\boldsymbol{\theta}) = -\frac{N}{2}\log\left(2\pi\sigma^2\right) + \log|\boldsymbol{I}_N - \lambda\boldsymbol{W}| - \frac{1}{2\sigma^2}\,\boldsymbol{u}'\boldsymbol{u}\,, \tag{3}$$

where $\boldsymbol{\theta} = (\boldsymbol{\beta}, \lambda, \sigma^2)$, $\boldsymbol{u} = \boldsymbol{y} - \lambda\boldsymbol{W}\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}$ and $\det(\partial\boldsymbol{u}/\partial\boldsymbol{y}) = |\boldsymbol{I}_N - \lambda\boldsymbol{W}|$ is the Jacobian. Using eigenvalues κ of matrix $\boldsymbol{W}$ derived by row standardization from a symmetric $\boldsymbol{C}$ matrix, [4] shows that condition $\lambda \in (\min(\kappa)^{-1},\ \max(\kappa)^{-1})$ must be fulfilled to ensure model stability. Under non-symmetric connectivity, eigenvalues of $\boldsymbol{W}$ can be complex and the stability restriction for λ is derived by [13] as $\lambda \in (1/\kappa_{min}, 1)$ where κ_{min} is the most negative purely real eigenvalue of $\boldsymbol{W}$. [1,12] provide a technical discussion of ML-based estimation approaches for various distributional assumptions and model generalizations.

Upon data (instrument) availability, instrumental variable regression (IVR) and generalized method of moments (GMM) approaches can also be used to estimate spatial models like (2). While the IVR/GMM estimators do not rely on distributional assumptions, their use is limited by actual availability of suitable instruments [1,8]. Another disadvantage of the IVR/GMM approach lies in potentially ending up with a λ parameter of Eq.(2) that is outside its parameter space [4]. This comes from the fact that – unlike the ML estimator – IVR/GMM approaches ignore the Jacobian term $|\boldsymbol{I}_N - \lambda\boldsymbol{W}|$ in (3).

The above-discussed estimators (ML/IVR/GMM) take spatial structure (matrix $\boldsymbol{W}$) as prior information – the underlying spatial dependency process is fixed ("known"), except for the λ-parameter. Nevertheless, any given spatial lag model of the type (8) can be repeatedly estimated under alternative spatial structure inputs (alternative $\boldsymbol{C}$ and $\boldsymbol{W}$ matrices) – either for stability verification [12] or with the purpose of finding a "best" model along some reasonable metric (e.g. mean squared error of prediction).

In contrast to ML/IVR/GMM approaches, Bayesian Markov Chain Monte Carlo (MCMC) methods can serve for two purposes simultaneously: model estimation and spatial structure selection [13]. Under the Bayesian approach, we can start with model (2) specification and establish some S alternative spatial structures $\boldsymbol{W}$ with equal prior probabilities of $1/S$. Say, different $\boldsymbol{C}$ connectivity matrices (1) are produced based on varying the τ thresholds. Next, we estimate parameters for each model (specifications differ in the $\boldsymbol{W}$-term only) using the Bayesian MCMC approach. Posterior model probabilities are used to select the spatial structure that reflects the data best [4].

Lam and Souza [10] propose a methodology that generates both $\boldsymbol{\beta}$ estimates and spatial structure estimates in a non-Bayesian context. Their estimation method combines IVR and a modified least absolute shrinkage and selection operator (LASSO). For detailed discussion, we start by generalizing the spatial lag element $\lambda \boldsymbol{W}\boldsymbol{y}$ and by transforming model (2) for panel data:

$$\boldsymbol{y}_t = \boldsymbol{\mu} + \left(\boldsymbol{A} + \sum_{s=1}^{S} \theta_s \boldsymbol{W}_s \right) \boldsymbol{y}_t + \boldsymbol{X}_t \boldsymbol{\beta} + \boldsymbol{\epsilon}_t, \quad t = 1, \ldots, T, \tag{4}$$

where $t = 1, \ldots, T$ is used to index time periods, $\boldsymbol{\mu}$ is a time-invariant vector of N individual effects and $s = 1, \ldots, S$ identifies each of the S distinct spatial structures considered (alternative setups can be based on distances, contiguity, transportation infrastructure, expert knowledge, etc.). The matrix $\boldsymbol{A}$ is an estimated $N \times N$ sparse "adjustment" element and there are S spatial dependency coefficients θ_s in model (4), instead of the single λ-coefficient in (2). Finally, $\boldsymbol{\epsilon}_t$ is the error term.

For estimation of (4), IVR is used to deal with endogeneity in the spatial lag element, as well as with potential endogeneity in the $\boldsymbol{X}$ regressors. The estimator proposed by [10] involves a LASSO penalty on the estimated elements of $\boldsymbol{A}$. Matrix $\boldsymbol{A}$ is assumed exogenous and its vectorized form $\boldsymbol{\xi} = (a_{11}, \ldots, a_{1,N}, a_{2,1}, \ldots, a_{N,N})'$ is part of the parameter vector that is estimated. LASSO penalty generates the desired sparsity of matrix $\boldsymbol{A}$.

For stability of model (4) estimation, $\sum_{s=1}^{S} \theta_s \leq 1$ is assumed along with row-wise restriction on the whole spatial structure matrix element of (4). The proposed IVR/LASSO estimator has several advantages: we can distinguish "realistic" and insignificant spatial setups $\boldsymbol{W}_s$ through statistical inference of their corresponding θ_s coefficients. Also, if the resulting $\hat{\boldsymbol{A}}$ (estimate of $\boldsymbol{A}$) is close to a zero matrix, then one can assume that proper (realistic, accurate) spatial structures $\boldsymbol{W}_s$ are used as prior information and no "adjustment" is necessary. Finally, the IVR/LASSO estimation can be performed even if prior spatial spec-

ification is absent, i.e. for $\boldsymbol{W}_s = \boldsymbol{0}$ for all s. In such case, $\hat{\boldsymbol{A}}$ is the estimated spatial structure and it acts as a spatial weight matrix.

Availability of large panel data is the main limiting factor for this estimator (along with spatial structure being invariant over extended time periods – i.e. for large T). Given the $N^2 - N$ estimated elements of matrix $\boldsymbol{A}$, the estimator requires both the T and N dimensions of the panel to be large [9]. Indeed, [10] provide an empirical example for their method that is based on stock-market data; which is not a typical use case for econometric spatial models.

Section 3 introduces a new panel data estimator: using the ML approach, the proposed method uses empirical sectoral data in a two-phase process that provides for spatial structure selection and subsequent main model estimation. Unlike the single-step LASSO estimator, the proposed estimator is suitable for "short" panels (N large, T small) or even cross sectional data.

2.3 Model Evaluation: Selection and Testing

Once we focus on evaluating spatial structures for a given econometric model, some important considerations arise:

1. Evaluation of alternative models can be approached using two paradigms: model selection and models testing – both concepts are widely used in empirical applications, yet they are not quite mutually compatible.
2. Regression models (2) that are based on different spatial weights $\boldsymbol{W}$ (i.e. on different resulting regressor vectors $\boldsymbol{Wy}$) are non-nested. Hence, the usual likelihood ratio test, Wald test, etc. cannot be formally applied if the testing paradigm is chosen [4,15].

Model selection and testing are two conceptually different tasks. In principle, model selection process approaches all models symmetrically, while the testing approach treats the null and alternative models differently. Model selection provides a definitive output – one model is selected. On the other hand, hypothesis testing does not seek such outcome: even if the null model is rejected, the alternative specification is not necessarily accepted. Also, the choice of null hypothesis is crucial for the test outcome.

The distinction between selection and testing is also empirically motivated: model selection is more appropriate for supporting decision making processes (implementation and evaluation of economic policies) while testing is used for inferential applications (e.g. if validity of a theoretically determined prediction is assessed).

In empirical applications, both approaches are relevant and worthy. [15] argues that current model selection methods are mostly based on statistical measures of model fit (sum of squared residuals, maximized log-likelihoods, information criteria, etc.). This brings model selection closer to hypothesis testing than the underlying conceptual differences would suggest. The empirical part of this contribution encompasses both evaluation methods: maximized log-likelihoods are used for model selection and the Vuong's test (6) is demonstrated.

Bayesian MCMC estimation methodology provides a nice example of model selection: estimation can be performed across different spatial setups and model selection is performed through maximizing posterior probabilities. Theoretically, such selection may lead to sub-optimal choices, as the approach might only find a local maximum among models (spatial setups) considered. [4] provides references to empirically-based arguments against practical limitations of this approach.

[15] lists penalized regression estimators as model selection methods. This is a reasonable view in the context of model (4) estimation by the IVR/LASSO approach – with no prior spatial information ($\boldsymbol{W}_s = \mathbf{0}$ for all s), the spatial weight matrix is estimated entirely. While this approach eliminates the risk of substantially sub-optimal spatial setup outcomes, the obvious disadvantage is the volume of observations necessary for estimation of all model parameters, which are on the order of N^2 [9].

For testing purposes, [15] provides several means of comparing non-nested specifications. However, most of the tests (J-test, JA-test, N-test, NT-test, etc.) are only suitable for OLS-estimated models. While such test may be extended to the IVR methodology proposed by [1], they are not applicable for ML-based estimators.

A feasible approach towards testing of non nested ML-estiated spatial models is provided by Vuong's test, which builds on the Kullback–Leibler information criterion (KLIC). Loosely speaking, KLIC measures the difference in maximized log-likelihoods between a misspecified model, say $f(y|\boldsymbol{Z}, \boldsymbol{\beta})$ and a true model $h(y|\boldsymbol{Z}, \boldsymbol{\alpha})$. For cross-sectional data, we can write

$$\text{KLIC} = E\left[\log h(y_i|\boldsymbol{z}_i, \boldsymbol{\alpha})|h \text{ is true}\right] - E\left[\log f(y_i|\boldsymbol{z}_i, \boldsymbol{\beta})|h \text{ is true}\right], \tag{5}$$

where $\boldsymbol{z}_i$ is a full set of regressors for the ith observation, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are model parameters. Even if the true specification $h(\cdot)$ is unknown, KLIC can be used to compare two alternative functions, say f_0 and f_1. [18] demonstrated that by taking the difference of KLIC of two given f_1 and f_0 functions, the true likelihood function h cancels out. For non nested models, Vuong's test statistic V for a "directional" test can be constructed as follows:

$$V = \frac{\sqrt{N}\left(\frac{1}{N}\sum_{i=1}^{N} m_i\right)}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(m_i - \overline{m})^2}} = \sqrt{N}\left(\overline{m}/s_m\right), \quad m_i = \log L_{i,1} - \log L_{i,0}, \tag{6}$$

where $L_{i,1}$ and $L_{i,0}$ are likelihood functions of f_1 and f_0 evaluated at a given observation i. Elements $\overline{m}$ and s_m refer to sample means and standard deviation of m_i. Under the null hypothesis of both tested models being equally good (i.e. equally distant from the true model h), V asymptotically follows standard Normal distribution. If f_1 is substantially better than f_0 (i.e. closer to h), V diverges and $\operatorname{plim} V = +\infty$ (and vice versa). Technical discussion on the test and distribution of V for nested and partially nested models is provided by [15,18].

3 Heuristic Enhancement of Spatial Information

Getting "some" spatial prior information for a regression model is often straightforward. However, assessing its validity and improving such spatial information is not an easy task. The approach proposed by [10] requires data (regressors and instruments) availability on the order of large (diverging) N and T. Unfortunately, such datasets are seldom available in regional economic analyses, which typically rely on short panel data.

For cross-sectional and short panel analyses, the search for a "better" spatial structure is necessarily heuristic, as the total number of distinct spatial structures equals $2^{(N^2-N)}$ for a time-invariant binary connectivity matrix $\boldsymbol{C}$ with zeros on the main diagonal. Authors concerned with the validity/robustness of spatial structures [5,14] often adopt a quasi-Bayesian approach by repeatedly estimating spatial models using systematically amended spatial weights. For distance-based neighborhoods, the τ parameter as in expression (1) can be iterated over some reasonable range. If neighborhood is defined along the kNN rule, then different k-values may be used.

Once several alternative models with varying spatial specifications are estimated, it is relatively easy to identify "underpeforming" spatial setups through examination of convenient statistics: minimized sums of squared errors, maximized log-likelihood statistics, etc. The same statistic can be used to select the "best" model as well. However, a number of similar spatial specifications may lead to models being statistically equivalent along the Vuong's test (6). Such result may be interpreted in terms of robustness and general validity of the spatial settings considered.

3.1 Motivation and Underlying Assumptions

Drawing neighborhood structures purely from aerial distances (including kNN-based neighbor classification) is rather restrictive. Room for improvement exists and the algorithm proposed next is designed to overcome the limitations of "hard" distance-based thresholds as in (1). The proposed heuristic approach is flexible, based on sectoral data and it can enhance spatial information used in regression models. Its validity and significance can be evaluated and tested as discussed in Sect. 2.3. Theoretical and empirical motivation of the proposed approach can be summarized as follows:

- Dissimilarities between regional distances and accessibility exist, along with local breaks in spatial patterns. Among the main factors for such behavior are the presence/absence of transport infrastructure, administrative barriers, common language, industrial clusters, etc.
- Direct measurement of the underlining factors would be tedious and prone to inconsistencies. However, by combining geo-referenced data of observed sectoral macroeconomic variables and statistical methods, we may extract spatial information from observed interactions among geographical units.

- The mainstream assumption of time-invariant spatial structures is made [4]. Actually, such assumption is more realistic for the short panel scenario as compared to the large T dimension that is necessary for the IVR/LASSO estimation of Eq. (4).
- Importantly, it is assumed that the true (yet unobserved) spatial pattern is reflected in observed co-dependencies among variables for a given sector and area (map). For example, as the empirical section of this paper focuses on labor market dynamics, the underlying spatial pattern is reflected in different sets of observed data: employment and unemployment indicators for diverse demographic groups (gender, age, education), job vacancies for different industrial sectors, etc. For other regional analyses, sectoral information can be gathered by analogy.
- The algorithm proposed avoids the risk of over-fitting by focusing on a limited number of potential "links" between regions: geographically very close units are always classified as neighbors and prominently distant units are treated as independent. The algorithm focuses on a relatively narrow "grey" area of distances around the τ threshold, where different pairs of regions may or may not be spatially dependent – depending of multiple underlying factors.

3.2 Combination of Distance-Based Information with Sectoral Data

The proposed method does not require "large" panels. Unlike the IVR/LASSO approach, this is essentially a two-step approach: In the first step, an adjustment spatial matrix $\boldsymbol{H}$ is constructed based on empirical sectoral "evidence" and combined with distance-based connectivity information. The second step is just an estimation of the main spatial regression model that is based on enhanced spatial connectivity matrix. If the above-discussed assumptions and sectoral data availability conditions are satisfied, construction of the enhanced connectivity matrix $\boldsymbol{C}_e$ and model estimation may be cast as follows:

1. For the dependent variable of a spatial regression model, we collect relevant sectoral variables for the time period and spatial units given. For example, we can use different variables that are related to the labor market.
2. Available geo-coded sectoral variables are centered, standardized (all data series have mean zero and variance one) and organized into a matrix $\boldsymbol{Z} = [z_{\ell,i}]$ that has a dimension $L \times N$. Each row $\ell = 1, \ldots, L$ of $\boldsymbol{Z}$ features a sectoral variable, each column $i = 1, \ldots, N$ corresponds to a given ith region.
3. Use sectoral data to calculate a correlation matrix $\boldsymbol{R} = [r_{ij}]$ for all pairs of spatial units $\{i, j\}$ – i.e. for all pairs of columns in $\boldsymbol{Z}$.
4. Individual elements of matrix $\boldsymbol{H}$ are constructed so that $h_{ij} = 1$ if $r_{ij} > \rho$ for some conveniently set ρ (say, $\rho = 0.8$).
   Note: This applies to instances where the dependent variable $\boldsymbol{y}$ in the main model exhibits positive spatial autocorrelation (e.g. when tested by Moran's I). In the somewhat unlikely case of negative spatial autocorrelation in $\boldsymbol{y}$, $\boldsymbol{H}$ would be based on "large" negative r_{ij} values. Combining positive and negative elements (albeit significant) in $\boldsymbol{H}$ would violate the process of constructing $\boldsymbol{W}$ by row-standardizing a connectivity matrix $\boldsymbol{C}_e$.

5. Next, matrix $\boldsymbol{C}_e$ is produced as follows:

$$\boldsymbol{C}_e = [c_{e,ij}] = \begin{cases} 0 & \text{if} \quad i = j\,, \\ 0 & \text{if} \quad d_{ij} > \tau^+\,, \\ h_{ij} c_{ij}^+ & \text{if} \quad \tau^+ \geq d_{ij} > \tau^-\,, \\ 1 & \text{if} \quad d_{ij} \leq \tau^- \quad \text{and} \quad i \neq j\,, \end{cases} \tag{7}$$

   where τ^- and τ^+ values describe the "grey" region of connectivity (say, the vicinity is given by $\tau \pm 15\%$). Here, h_{ij} elements are used to enhance spatial information. At distances below τ^-, all elements are neighbors. With $d_{ij} > \tau^+$, regions i and j are always independent.
6. Step 5 may be performed repeatedly, for a conveniently chosen sequence of τ values. This way, alternative spatial setups are generated for subsequent model estimation, selection and testing.
7. The main spatial regression model is estimated using an enhanced spatial structure. Before estimating spatial model using an enhanced spatial information, matrix $\boldsymbol{C}_e$ is transformed into a spatial weight matrix $\boldsymbol{W}$ by row-standardization (distinct notation for such $\boldsymbol{W}$ matrix is not necessary for this article).

The resulting $\boldsymbol{C}_e$ matrix from (7) is symmetric, potentially missing some irregularities in spatial patterns. Nevertheless, the algorithm outlined points towards a general concept of empirically enhanced neighborhood structures. Diverse amendments to the algorithm are possible. The random forest and similar machine learning algorithms can be used instead of the pairwise correlations to generate non-symmetric $\boldsymbol{H}$ matrices [7].

Describing spatial structures using $\boldsymbol{H}$ alone is somewhat problematic: it would typically lead to over-fitting and would generate long-distance "spurious" connections that contradict the underlying assumptions of spatial analysis, rooted in proximity-based interactions.

3.3 Enhanced Spatial Structures: Robustness and Stability Aspects

Enhancing neighborhood structure information through sectoral data may bring significant improvement, yet arbitrary aspects are not fully eliminated. For the algorithm described and for expression (7), the choice of sectoral variables may affect the outcome, as well as the definition of "grey" area interval $\langle \tau^-, \tau^+ \rangle$.

The choice of sectoral data is relatively straightforward: statistical offices and central authorities typically provide datasets in a hierarchically structured manner, which facilitates both choosing and disclosing data sources. This does not eliminate the heuristic nature of choosing data, yet provides transparency and reproducibility for subsequent verification and further applications.

Finding a proper $\langle \tau^-, \tau^+ \rangle$ interval may seem more fuzzy. However, spatial regression models under scrutiny can be repeatedly estimated while small & systematic amendments to neighborhood definitions are imposed [6]. For example, if $\tau \pm 15\%$ is used to generate the sought interval, we may repeatedly perform the

(7) neighborhood construction, based on small increments of the center-point τ distance [14].

Once enough consistently made changes $\langle \tau^-, \tau^+ \rangle$ are evaluated, we can assess model features (maximized ML values, predictive properties, etc.) and search for the "best" spatial setup (ceteris paribus). Typically, we would find underperforming spatial setups as well as clusters of models with stable properties (robust spatial settings).

4 Empirical Illustration Based on Labor Market Data

This section provides a demonstration of the potential benefits brought by enhancing spatial information (spatial structure) through relevant sectoral data. Empirical illustration is based on regional (NUTS2) labor market data, 2015–2019 annual observations covering the following EU countries: Austria, Belgium, Czechia, Denmark, Hungary, Luxembourg, the Netherlands, Poland, Slovenia and Slovakia. The cross-sectional sample is illustrative, the time domain is chosen to avoid Covid19-related disturbances in the data generating process. Enhanced spatial information is compared to traditional approaches.

Labor market dynamic processes are estimated through a spatial lag panel model [4], with unemployment rates cast as a linear function of GDP growth (in terms of log-transformed GDP per capita), individual (regional) effects, state-level (NUTS0) hierarchical effects and spatial autoregressive effects in the dependent variable. The panel model can be formulated as:

$$\boldsymbol{y} = \lambda \left(\boldsymbol{I}_T \otimes \boldsymbol{W} \right) \boldsymbol{y} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}\,, \tag{8a}$$

$$\boldsymbol{u} = \left(\boldsymbol{\iota}_T \otimes \boldsymbol{I}_N \right) \boldsymbol{\mu} + \boldsymbol{\varepsilon}\,, \tag{8b}$$

$$\boldsymbol{x}_{it}\boldsymbol{\beta} = \alpha + \beta \, \log \left(GDP_{i,t-1} \right) + NUTS0_i' \boldsymbol{\gamma}\,, \tag{8c}$$

where $\boldsymbol{y}$ is the dependent variable. Each y_{it} element of this $(NT \times 1)$ vector represents one ith region ($N = 110$, i.e. there are 110 NUTS2 regions in the panel) observed at time t and $\boldsymbol{W}$ is the spatial weights matrix. Dimensions of identity matrices $\boldsymbol{I}_T$ and $\boldsymbol{I}_N$ are apparent from their subscripts and $\boldsymbol{\iota}_T$ is a $(T \times 1)$ vector of ones. The symbol $\otimes$ represents Kronecker product. Equation (8b) shows the composition of random terms u_{it}, cast as the sum of unobserved individual effects μ_i (note that spatial heterogeneity is assumed time-invariant which is reflected in the subscript) and the idiosyncratic terms ε_{it}. Equation (8c) outlines the regressors used in matrix $\boldsymbol{X}$ – the LHS element $\boldsymbol{x}_{it}$ denotes a row of $\boldsymbol{X}$-matrix, i.e. a set of regressors for a given observation. Vector $\boldsymbol{\beta} = (\alpha, \beta, \gamma_1, \ldots, \gamma_9)'$ and the coefficient λ are parameters of the model, estimated by ML. Generalized panel data versions of the likelihood function (3) are provided by [14], along with detailed technical discussion on estimation and statistical inference. Based on parameter estimates and the spatial weight matrix $\boldsymbol{W}$ given, marginal effects are calculated [4].

Data used for estimation of model (8) – both macroeconomic data and geographical information – are drawn from Eurostat. For the sake of reproducibility, dataset identification is provided: unemployment rate (dependent variable)

comes from "lfst_r_lfu3rt" and the log-transformed GDP per capita is based on "nama_10r_2gdp". Vector $NUTS0_i'$ is made of nine dummy variables, identifying how individual NUTS2 regions belong to their corresponding states. Hence, besides regional (individual) effects addressed by the μ_i elements of (8), state-level hierarchical structure is accounted for as well. In $NUTS0_i'$, Austria is the reference (omitted) factor and γ_j parameters control for state-specific behaviour (heterogeneity) of unemployment.

Individual elements h_{ij} of the empirically determined matrix $\boldsymbol{H}$ are obtained through correlation analysis. Matrix $\boldsymbol{C}_e$ as in (7) was estimated using sectoral labor market data. For the NUTS regions and time periods used for estimation of the main model (8), sectoral variables for different demographic groups and employment & unemployment classifications are utilized. Specifically, Eurostat datasets "lfst_r_lfu3rt", "tgs00007", "tgs00054" and "tgs00102" were used.

For demonstration purposes, unemployment dynamics as in model (8) are estimated using three different approaches to spatial structure construction, reflected in the $\boldsymbol{W}$ matrix. First, a purely distance-based spatial structure is set along expression (1). Second, enhanced spatial structure following expression (7) is used. Finally, contiguity-based spatial structure is used for comparison. For distance-based spatial weights, model (8) is estimated repeatedly over a series of τ values – and corresponding $\tau \pm 15\%$ ranges in enhanced spatial setups. Distances from $\tau = 200$ km to $\tau = 800$ km (with 10-km threshold increments) are used to evaluate model stability and statistical properties under varying spatial prior information.

Repeated estimations of model (8) – a total of 61 different models for each of the spatial definitions (1) and (7) – generate large volumes of parameters, marginal effects, standard errors and corresponding statistics. To compare and evaluate model estimates across varying spatial settings, Fig. 2 is provided, showing maximized log-likelihood values, spatial dependency parameter λ and marginal effects (direct & spillover) of GDP. The x-axes of individual elements of Fig. 2 show the corresponding τ thresholds.

From the top-left plot of Fig. 2, we may compare the maximized log-likelihood (LL) values of estimated models: red line represents models with purely distance-based $\boldsymbol{W}$ matrices and blue line represents enhanced spatial structures using expression (7). Based on the LL information, the "best" spatial structure is inferred from an enhanced spatial structure that is generated for $\tau = 220$ km. Another distinct local maximum is apparent at $\tau = 330$. Both thresholds of interest are highlighted by vertical dotted lines in the plot. In terms of LL values, enhanced spatial structure (blue line) is superior to the purely distance-based setup. This conclusion holds over the whole range of thresholds that have been evaluated. Contiguity-based estimation is not shown in Fig. 2 as it cannot be reasonably plotted against the x-axis. Nevertheless, relevant estimation output is provided in Table 1.

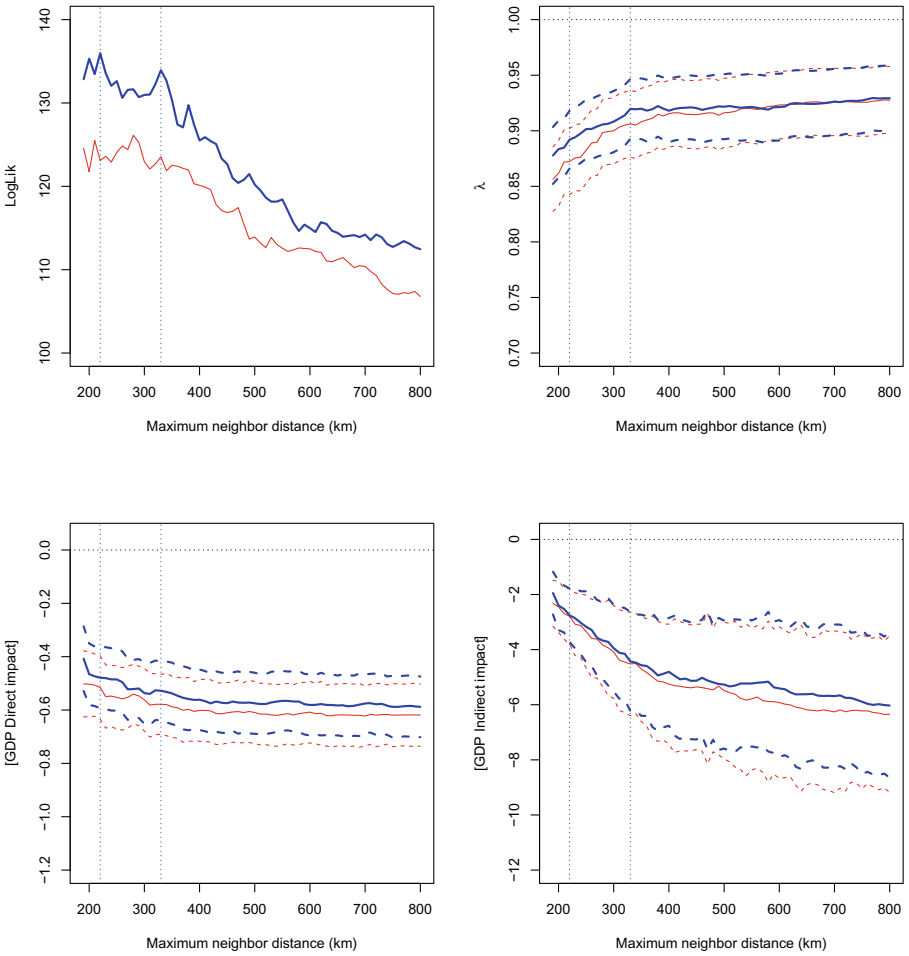


Fig. 2. Model evaluation & impact estimates under varying spatial structures. Enhanced spatial information (blue) is compared to distance-based spatial structures (red) (Color figure online)

Direct comparison of the LL values is the main approach to evaluate and select different spatial structures used in regression models [14]. However, as Vuong's test (6) is applied, we may conclude that the enhanced spatial structure at $\tau = 220$ km is significantly closer to the true specification at the 5% significance level, when compared to estimation based on distances only (same τ) and to contiguity-based spatial weights, with Vuong's statistic values at $+2.82$ and $+2.23$ respectively.

The upper-right element of Fig. 2 shows the estimated spatial dependency parameter λ, along with corresponding $\pm$ 1 standard errors (dashed lines). Similarly, direct and spillover effects of GDP are shown in the two bottom elements

of Fig. 2. The $\boldsymbol{\gamma}$-parameters and corresponding marginal effects from Eq. (8c) are omitted from Fig. 2 and Table 1. The hierarchical structure of individual and state-level effects has to be controlled for to secure proper labor market dynamics estimation. However, individual state-level effects bear limited value in terms of direct interpretation. All omitted estimation outputs are available from the author upon request, along with data and R-codes.

Table 1. Comparison of model (8) estimates for alternative spatial structures

Impact/λ	Estimate	Std. Error (simulated)	z-value (simulated)	Pr ($>\|z\|$) (simulated)
Enhanced spatial structure used, $\tau = 220$ km:				
GDP_Direct_Imp	–0.479	0.113	–4.116	0.000
GDP_Indirect_Imp	–2.755	0.972	-2.895	0.004
λ	0.892	0.026	34.339	0.000
Log likelihood (*LL*) 135.973				
Enhanced spatial structure used, $\tau = 330$ km:				
GDP_Direct_Imp	–0.528	0.110	–4.813	0.000
GDP_Indirect_Imp	–4.424	1.798	–2.632	0.008
λ	0.920	0.027	34.232	0.000
Log likelihood (*LL*) 133.927				
Purely distance-based spatial structure, $\tau = 220$ km:				
GDP_Direct_Imp	–0.516	0.119	–4.287	0.000
GDP_Indirect_Imp	–2.788	1.009	–2.844	0.004
λ	0.872	0.030	29.351	0.000
Log likelihood (*LL*) 123.090				
Purely distance-based spatial structure, $\tau = 330$ km:				
GDP_Direct_Imp	–0.578	0.112	–5.185	0.000
GDP_Indirect_Imp	–4.518	1.880	–2.589	0.010
λ	0.906	0.030	30.273	0.000
Log likelihood (*LL*) 123.536				
Contiguity-based spatial structure				
GDP_Direct_Imp	–0.575	0.119	–4.853	0.000
GDP_Indirect_Imp	–0.920	0.256	–3.675	0.000
λ	0.662	0.040	16.630	0.000
Log likelihood (*LL*) 54.293				

As shown in Fig. 2, marginal effects estimated using two adjacent τ thresholds are very alike and cannot be distinguished at the 5 % significance level. This result is not surprising as it corresponds with the conclusions drawn by [12]: we cannot expect materially different marginal effects from two spatial models that are estimated using strongly correlated spatial settings (and are identical in their $\boldsymbol{X}$ matrix regressors).

Nevertheless, two important implication can be made based on Fig. 2: The use of varying maximum distance thresholds over a relatively extended range of

feasible τ-values allows us to find efficient spatial structures for model estimation. Along the LL criteria, enhanced spatial specification at $\tau = 220$ km threshold is selected as the "final" specification. Using Vuong's test (6), enhanced spatial specification provides significant improvement over both the contiguity-based and purely distance-based spatial structures.

Table 1 provides detailed information on model (8) estimates generated using the empirically determined $\tau = 220$ km and $\tau = 330$ km thresholds. Contiguity-based spatial structure is provided for comparison. All estimated λ-coefficients and impacts are significant at the 5% level. For all distance-based spatial settings shown in Table 1, the spillover effects of GDP are almost an order of magnitude more prominent when compared to direct effects. This strongly supports the validity of spatial methods for this type of analysis. Negative values of the estimated impacts (direct and spillover) follow from macroeconomic theory, which implies inverse relationship between GDP and unemployment dynamics. Estimates of the λ-coefficients suggest a strong spatial autocorrelation process in the dependent variable and their simulated z-scores are very high – this result also reflects the relative simplicity of the illustrative model (8).

5 Conclusions

A heuristic algorithm for generating enhanced spatial information for spatial econometric models is presented. The proposed approach draws spatial information from sectoral economic data. The algorithm presented can be used for a wide range of economic and non-economic applications. The enhancement of spatial structures is demonstrated through a labor market-based empirical analysis, using panel data at the NUTS2 level for 10 selected EU countries with annual observations covering the period 2015 to 2019.

Overall, it should be noted that the use of additional (e.g. sectoral) information in itself does not guarantee an improved performance in a given empirical application. However, the efficiency of the proposed method can be easily tested for any given dataset/usecase by means of the Vuong's test.

Empirically, enhanced spatial information show significant improvement over models estimated using neighborhood information based purely on distances or contiguity. However, this approach has some inherent limitations to its usage, especially if sectoral data are scarce. Also, some general stability assumptions apply to the underlying data generating processes.

Spatial analyses and spatial econometric models always face some level of uncertainty with respect to the true yet unobservable spatial structure (neighborhood definition). This article provides an intuitive, relatively simple and empirically sound algorithm that helps mitigate such uncertainty.

Acknowledgement. Supported by the grant No. IGA F4/38/2022, Faculty of Informatics and Statistics, Prague University of Economics and Business.

References

1. Anselin, L.: Spatial Econometrics: Methods and Models. Kluwer, Dordrecht (1988)
2. Baltagi, B.H.: Econometric Analysis of Panel Data. Wiley, New York (2005)
3. Croissant, Y., Millo, P.: Panel data econometrics in R: the PLM package. J. Stat. Softw. **27**(2), 1–43 (2008). https://doi.org/10.18637/jss.v027.i02
4. Elhorst, J.P.: Spatial Econometrics: From Cross-Sectional Data to Spatial Panels. Springer-Briefs in Regional Science, Springer, Berlin (2014). https://doi.org/10.1007/978-3-642-40340-8
5. Formánek, T.: Semiparametric spatio-temporal analysis of regional GDP growth with respect to renewable energy consumption levels. Appl. Stoch. Model. Bus. Ind. **36**(1), 145–158 (2020). https://doi.org/10.1002/asmb.2445
6. Formánek, T.: Spatially augmented analysis of macroeconomic convergence with application to the Czech republic and its neighbors. In: Silhavy, R., Silhavy, P., Prokopova, Z. (eds.) CoMeSySo 2017. AISC, vol. 662, pp. 1–12. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-67621-0_1
7. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer, New York (2001). https://doi.org/10.1007/978-0-387-21606-5
8. Kelejian, H.H., Prucha, I.R.: A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. J. Real Estate Finan. Econ. **17**(1), 99–121 (1998). https://doi.org/10.1023/A:1007707430416
9. Kim, Y., Hao, J., Mallavarapu, T., Park, J., Kang, M.: Hi-LASSO: high-dimensional LASSO. IEEE Access. **7**, 44562–44573 (2019)
10. Lam, C., Souza, P.C.L.: Estimation and selection of spatial weight matrix in a spatial lag model. J. Bus. Econ. Stat. **38**(3), 693–710 (2020). https://doi.org/10.1080/07350015.2019.1569526
11. Lee, L., Yu, J.: Some recent developments in spatial panel data models. Reg. Sci. Urban Econ. **40**(5), 255–271 (2010). https://doi.org/10.1016/j.regsciurbeco.2009.09.002
12. LeSage, J.P., Pace, R.K.: The biggest myth in spatial econometrics. Econometrics. **2**(4), 217–249 (2014). https://doi.org/10.3390/econometrics2040217
13. LeSage, J.P., Pace, R.K.: Introduction to Spatial Econometrics. CRC Press, Boca Raton (2009)
14. Millo, G., Piras, G.: splm: Spatial panel data models in R. J. Stat. Softw. **47**(1), 1–38 (2012). https://doi.org/10.18637/jss.v047.i01
15. Pesaran, M.H.: Time Series and Panel Data Econometrics. Oxford University Press, Oxford (2015)
16. Lovelace, R., Nowosad, J., Muenchow, J.: Book on Geographic Data with R. CRC Press, Boca Raton (2019)
17. Oshan, T.M.: The spatial structure debate in spatial interaction modeling: 50 years on. Prog. Hum. Geogr. **45**(5), 925–950 (2020). https://doi.org/10.1177/0309132520968134
18. Vuong, Q.H.: Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica **57**(2), 307–333 (1989). https://doi.org/10.2307/1912557

Use of Neighborhood-Based Bridge Node Centrality Tuple for Preferential Vaccination of Nodes to Reduce the Number of Infected Nodes in a Complex Real-World Network

Natarajan Meghanathan(✉), Kapri Burden, and Miah Robinson

Department of Electrical and Computer Engineering and Computer Science, Jackson State University, 1400 Lynch Street, Jackson, MS 39217, USA
natarajan.meghanathan@jsums.edu

Abstract. We investigate the use of the Neighborhood-based Bridge Node Centrality (NBNC) tuple to choose nodes for preferential vaccination so that such vaccinated nodes could provide herd immunity and reduce the spreading rate of infections in a complex real-world network. The NBNC tuple ranks nodes on the basis of the extent they serve as bridge nodes in a network. A node is a bridge node, if when removed its neighbors are either disconnected or at least sparsely connected. We hypothesize that preferentially vaccinating such bridge nodes would block an infection spread from a neighbor of the bridge node to an another neighbor that are otherwise not reachable to each other. We evaluate the effectiveness of using NBNC vis-a-vis degree centrality for preferential vaccination to reduce the spread of infections by conducting simulations of the spread of infections per the SIS (Susceptible-Infected-Susceptible) model on a collection of 10 complex real-world social networks.

Keywords: Vaccination · Bridge Node · Centrality Tuple · SIS Model · Simulations · Infection Spread

1 Introduction

With the COVID-19 pandemic creating havoc worldwide for the last few years, and the necessity of vaccinating people with boosters to protect them from getting infected with variants of the virus, it becomes imperative to identify effective strategies to preferentially vaccinate people in a social network or a community so that such vaccinated people could provide herd immunity (i.e., block the infection from spreading to the non-vaccinated people) and thereby the average number of infected people per network or community is eventually reduced. Network Science provides solution to the above problem through the notion of "Centrality" metrics. Centrality metrics quantify the topological importance of a node in the network [14]. Degree centrality (DEG) is the most computationally-light metric that has been often used (e.g., [15]; most recent use reported in [1]) to rank the nodes for preferential vaccination under the premise that nodes with several neighbors are more likely to spread the infection from one neighbor to another neighbor and vaccinating high DEG nodes could reduce the overall number of infected nodes.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 27–36, 2024.
https://doi.org/10.1007/978-3-031-53552-9_3

In a recent work [11], the author proposed the notion of neighborhood-based bridge node centrality (NBNC) tuple to quantify and ranked based on the extent to which nodes play the role of bridge nodes. A node is a bridge node [11, 16] if when removed the neighbors of the node either get disconnected or are more likely to be sparsely connected (i.e., have to reach through a longer path of length much greater than 2). Our hypothesis in this paper is that bridge nodes (rather than high DEG nodes) would be a better choice for preferentially vaccinating nodes to attain herd immunity and reduce the overall number of infected nodes. Our hypothesis stems from the criteria used to rank the nodes as bridge nodes per the NBNC tuple.

The NBNC tuple of a node has three entries (the number of components in the neighborhood graph of the node, the algebraic connectivity ratio of the neighborhood graph of the node, the number of vertices in the neighborhood graph, which is also the degree of the node), A node with a higher degree may still have its neighbors connected (either directly or through a multi-hop path) if the neighbor node is removed from the network. On the other hand, if a node has several components in its neighborhood graph, it is more likely that any two neighbors of the node are disconnected or connected through a long multi-hop path when the node is removed from the network. Note that DEG is also the last of the three entries in the tuples for NBNC; if two nodes cannot be differentiated based on the first two entries in the NBNC tuple, then the degree of the nodes could be used to break the tie. Hence, NBNC (whose tuple formulation includes DEG as the last entry) is a more comprehensive centrality tuple (compared to a scalar DEG centrality metric) to rank the bridge nodes for preferential vaccination.

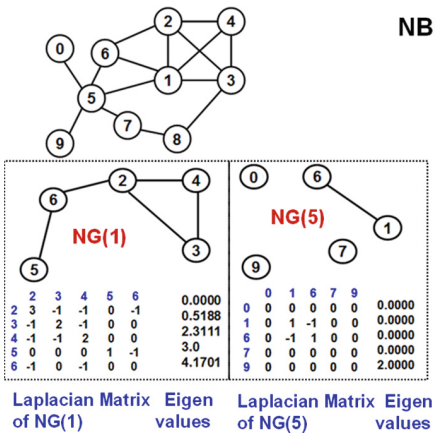
We use the SIS (susceptible-infected-susceptible) model [17], one of the widely used models for simulating the spread of infections. The R_0 (basic reproduction number) for a disease [17] is defined as the number of infections an infected individual could cause to a completely susceptible population. If R_0 for a disease is greater than 1, then one infected individual could lead to more than one newly infected individual and the disease would keep spreading. On the other hand, if R_0 for a disease gets less than 1, the chances of one infected individual leading to another infected individual gets lower and the disease will eventually die down. One way to reduce the R_0 for a disease is to preferentially vaccinate some nodes so that even if these nodes are exposed to the infected individuals, they would not get infected; as a result, if susceptible (non-vaccinated) individuals are exposed to the vaccinated individuals (but not to the infected individuals), the infection would not spread. At the same time, the infected individuals would also eventually recover and the disease would eventually die down. Hence, the motivation for this paper is to explore the use of NBNC for preferentially vaccinating nodes in the presence of a disease spread simulation (in rounds) per the SIS model and evaluate whether this leads to a lower value for the average number of infected nodes per round of the simulation (compared to the strategy of vaccinating nodes based on node degree, the currently preferred strategy).

The rest of the paper is organized as follows: Sect. 2 presents the notion of the NBNC tuple. Section 3 presents a simulation procedure for running the SIS model on a graph in the presence of a certain fraction of vaccinated nodes per the NBNC tuple vs. DEG centrality. Section 4 presents the results of the SIS simulations conducted for a suite of 10 complex real-world networks and compares the average number of infected individuals per round incurred with the NBNC tuple vs. DEG centrality-based

vaccinations. Section 5 presents related work in the literature. Section 6 concludes the paper and presents plans for future work. Throughout the paper, the terms 'node' and 'vertex', 'link' and 'edge', 'network' and 'graph' are used interchangeably. They mean the same.

2 Neighborhood-Based Bridge Node Centrality (NBNC) Tuple

The NBNC tuple of a node v is determined based on the neighborhood graph (NG) of the node. The neighborhood graph (NG) of a node v comprises of just the neighbor nodes of v and the edges connecting these neighbor nodes. The NBNC tuple for a node v has three entries, represented in this order: $NBNC(v) = [NG(v)^{\#comp}, NG(v)^{ACR}, |NG(v)|]$. $NG(v)^{\#comp}$ is the number of components in the neighborhood graph of node v. If the neighbors of node v are connected (reachable to each other) even after node v is removed from the network, then $NG(v)^{\#comp} = 1$; otherwise $NG(v)^{\#comp} > 1$. The second entry $NG(v)^{ACR}$ is the algebraic connectivity ratio of the neighborhood graph; the algebraic connectivity [18] of the neighborhood graph is the second Eigenvalue of the Laplacian matrix [19] of the neighborhood graph. If $NG(v)^{\#comp} > 1$, then the neighborhood graph is not connected and $NG(v)^{ACR}$ is 0; otherwise, $NG(v)^{ACR}$ is computed by dividing the algebraic connectivity by the number of nodes in the neighborhood graph (which also corresponds to the degree of node v). The third entry is the number of vertices in the neighborhood graph of the node (the node degree itself).



$$\text{NBNC(i)} = \left(NG_i^{\#comp}, NG_i^{ACR}, |NG_i|\right)$$

Node ID	*NBNC* Tuple
0	(1, 1.0, 1)
1	(1, 0.1038, 5)
2	(1, 0.25, 4)
3	(2, 0.0, 4)
4	(1, 1.0, 3)
5	(4, 0.0, 5)
6	(1, 0.3333, 3)
7	(2, 0.0, 2)
8	(2, 0.0, 2)
9	(1, 1.0, 1)

Fig. 1. Example Graph and the NBNC Tuples of its Vertices

Figure 1 presents a toy example graph that will be used in this section as well as in the next section. Figure 1 also presents the neighborhood graphs for some of the vertices of the graph; their Laplacian matrix and the Eigenvalues of the Laplacian matrix. The entries in a Laplacian matrix for a neighborhood graph are as follows: The diagonals indicate the number of neighbors for the nodes in the neighborhood graph; an entry (u, v) is -1 if there is an edge between u and v in the neighborhood graph; otherwise, the entry is a 0.

The following ranking criteria is used to rank nodes on the basis of the extent they play the role of bridge nodes. Nodes with more components in their neighborhood graph are ranked higher. If two nodes have the same number of components in their neighborhood graph, then the tie is broken in favor of the node with sparsely connected neighborhood (i.e., the node with the lower algebraic connectivity ratio is ranked higher). If two nodes cannot be differentiated based on the first two entries, then the tie is broken in favor of larger node degree. If two nodes cannot be differentiated based on all the three entries of their NBNC tuples, then they are equally ranked.

3 Procedure to Simulate the SIS Model

We use the SIS (Susceptible-Infected-Susceptible) model [17] to evaluate the effectiveness of NBNC tuple-based vaccination of nodes vis-a-vis the DEG centrality-based vaccination. Under this model, a node remains either in susceptible state or infected state. A susceptible node could get infected with a probability β and an infected node could recover with a probability μ. Once recovered, the node enters the susceptible state. We start with a graph of 'N' nodes and decide to vaccinate a certain fraction (λ) of nodes. The nodes that constitute the λ fraction of nodes to be vaccinated are chosen based on their ranking with respect to either the NBNC tuple or the DEG centrality values.

As part of initialization, we generate a random number for each node and set the node to either susceptible state (if the random number generated is greater than β) or infected state (if the random number generated is less than or equal to β). The simulation proceeds in rounds. Each round has two phases, executed in this sequence: (phase-i) Infected nodes changing their state to Susceptible with a probability μ; (phase-ii) Nodes that still remain infected after phase-i will infect their susceptible neighbors (i.e., neighbor nodes that are neither vaccinated nor infected) with a probability β. We count the number of nodes that are in the Infected state after phase-ii and add it to the total number of infected nodes across all the rounds. The simulation is stopped in one of these two ways: (1) A round of simulation is run if at least one node stays Infected after phase-i of the round; otherwise, the simulation stops. (2) We run the simulations for a maximum number of rounds and then stop. After the simulation has stopped, we determine the average number of infected nodes per round by dividing the total number of infected nodes across all the rounds by the total number of rounds the simulation was run.

Simulation of a Sample Round. Figure 2 presents the execution of a sample round of the simulation. Let the parameters β and μ be 0.5 each. Figure 2-(a) shows the graph at the beginning of the round (i.e., before phase-i is executed). To execute phase-i, we generate a random number (in the range of 0…1) at each infected node: if the random number comes out to be less than or equal to μ, the infected node is considered to have recovered and moves from Infected state to Susceptible state. In Fig. 2-(a) and 2-(b), we notice nodes 0, 6 and 9 (with a random number less than 0.5 for each of them) recover and become susceptible. During phase-ii, for each infected node, we generate a random number for the link with its susceptible neighbors: if the random number comes out to be less than or equal to β, then that susceptible neighbor node is considered to have become infected. Both the neighbors of node 8 are infected; there is a 50% chance that node 8 could get infected due to one of them and it happens so due to node 7. On the

other hand, among the three neighbors of node 6, only one of them (node 2) is in the Infected state (and the other two nodes are susceptible and vaccinated; so no infection could occur due to these two nodes); the chances of node 6 becoming infected is only 1/3 and the random number generated for the link 2–6 is greater than $\beta = 0.5$.

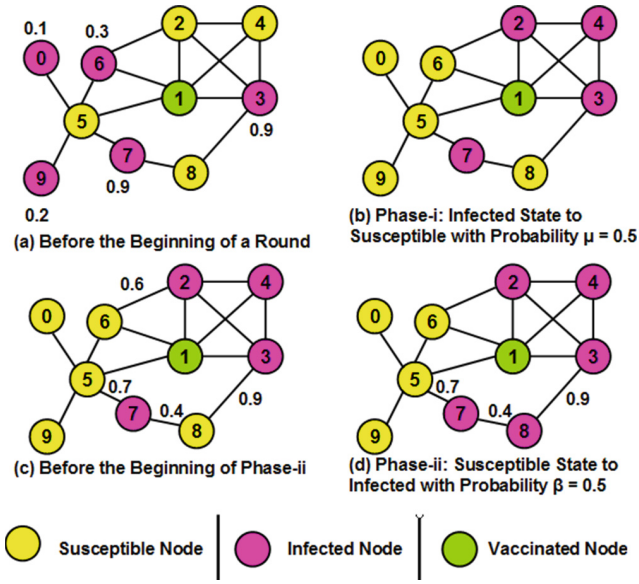


Fig. 2. A Sample Round of Simulation: Phase-i and Phase-ii

A susceptible node that is surrounded only by infected neighbor nodes manages to stay in Susceptible state only if the random numbers generated for each of those links with the infected neighbor nodes is greater than β. This is where vaccination comes helpful. If a vaccinated node is the only neighbor for a susceptible node, then the latter will stay forever as susceptible and will never become infected. Even if a vaccinated node is one of the few neighbors of a susceptible node, the susceptible node is less likely to become infected during any round. On the other hand, if a susceptible node is surrounded by several infected neighbor nodes, the node is likely to become infected during any round.

4 SIS Simulations for Real-World Networks

We conducted simulations of the infection spread (per the SIS model) for a collection of 10 real-world networks. Table 1 lists the networks and their IDs used in Figs. 3 and 4. The simulation parameters for the SIS model are: (β) - probability with which a susceptible node becomes an infected node during any round; (μ) - probability with which an infected node gets cured and returns to the Susceptible state during any round. We also vaccinate λ fraction of nodes for any simulation and the nodes to be vaccinated

are chosen based on their ranking with respect to either the NBNC tuple or DEG (degree centrality).

Table 1. Real-World Networks used in the Simulations

Net-ID	Real-World Network Graph	# Nodes
Net-1	Taro Exchange Network	22
Net-2	Sawmill Striker's Network	24
Net-3	Karate Network	34
Net-4	Teenage Women Friends Network	50
Net-5	Lazega Law Firm Network	71
Net-6	Copperfield Network	87
Net-7	US Football 2001 Network	105
Net-8	Anna Karnenina Network	138
Net-9	Jazz Band Network	198
Net-10	CKM Physicians Network	246

The operating conditions of a simulation on a particular real-world network are each possible combination of the above different values of the parameters β {0.3, 0.5, 0.7}, μ {0.25, 0.5} and λ {0.05, 0.10, 0.15, 0.20, 0.30} as well as the vaccination strategy (NBNC or DEG-based), leading to a total 3 * 2 * 5 * 2 = 60 operating conditions. We conduct 50 trials of the simulations for each operating condition and each simulation is run for a maximum of 20 rounds (a simulation stops prematurely before 20 rounds, if no nodes are infected for the next round) and we measure the average number of infected nodes per round (across all the rounds and trials) for each of the 10 real-world networks under each of the 60 operating conditions. We compute two metrics based on the average number of infected nodes per round for each network: (1) The average fraction of infected nodes per round for NBNC vs. DEG: computed as the ratio of the average number of infected nodes per round and the number of nodes for the network and (2) The ratio of the average fraction of infected nodes based on DEG as the node selection strategy for vaccination and the average fraction of infected nodes based on NBNC as the node selection strategy for vaccination.

Figure 3 presents a heat map-based visualization of the numbers/results for the average fraction of infected nodes per round of the simulations for each of the 10 real-world networks and with respect to either NBNC or DEG as the node selection strategy for vaccination. The color coding in Fig. 3 is based on the premise that lower the value for the average fraction of infected nodes per round, the more effective is the particular centrality-based node selection strategy for vaccination: accordingly, the colors of the cells featuring these raw values range from red (larger value for the average fraction of infected nodes per round) to green (lower value for the average fraction of infected nodes per round).

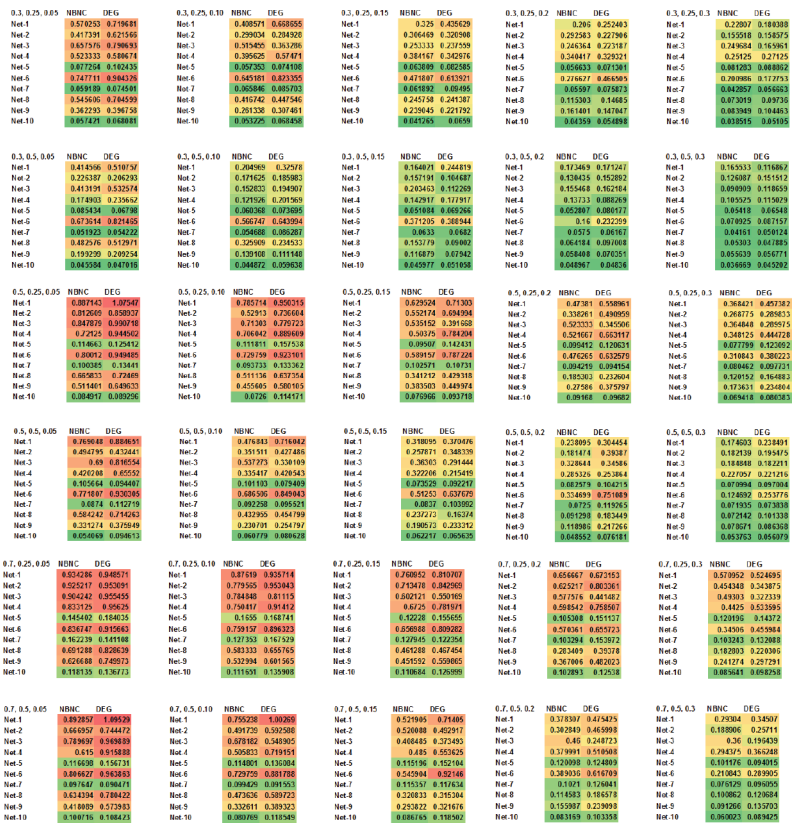


Fig. 3. Visualization of the Average Fraction of Infected Nodes Incurred per Round with respect to NBNC and DEG SIS Simulations

As expected for both NBNC and DEG, the transition in the colors of the cells from red to green (in Fig. 3) occurs gradually as we move across the scenarios exposing more nodes to infection (scenarios where $\beta \geq \mu$ and lower values of λ) to scenarios wherein nodes would not or are less likely to get infected (scenarios with larger values of λ and $\beta < \mu$). Each row in Fig. 3 corresponds to a particular combination of β and μ, with the λ values increased from 0.05 to 0.30; the colors of the cells accordingly transition from red to green as we increase λ for a given combination of β and μ. For most of the operating conditions and real-world networks, the average fraction of infected nodes with DEG is noticeably greater than that the average fraction of infected nodes with NBNC as the node selection strategy for vaccination. As a result, any particular record in Fig. 3 (for a particular network, for a given combination of λ, β and μ) is more likely to have the cell for NBNC to be less red (or more yellow or more green) compared to the cell for DEG.

Figure 4 presents the ratio of the average number of infected nodes (DEG) and average number of infected nodes (NBNC) for the 10 networks when the simulations are run for a particular combination of β and μ, with the λ values increased from 0.05 to 0.30. We observe the ratio to be heavily distributed above the line for 1.0, confirming our

claim that NBNC would be more appropriate to choose nodes for vaccination vis-a-vis DEG. We observe fewer ratios to be less than 1.0 for scenarios in which more nodes are exposed for infection (i.e., when $\beta \geq \mu$). The median of the ratio values are typically in the range of 1.15 to 1.20 for most of the operating conditions.

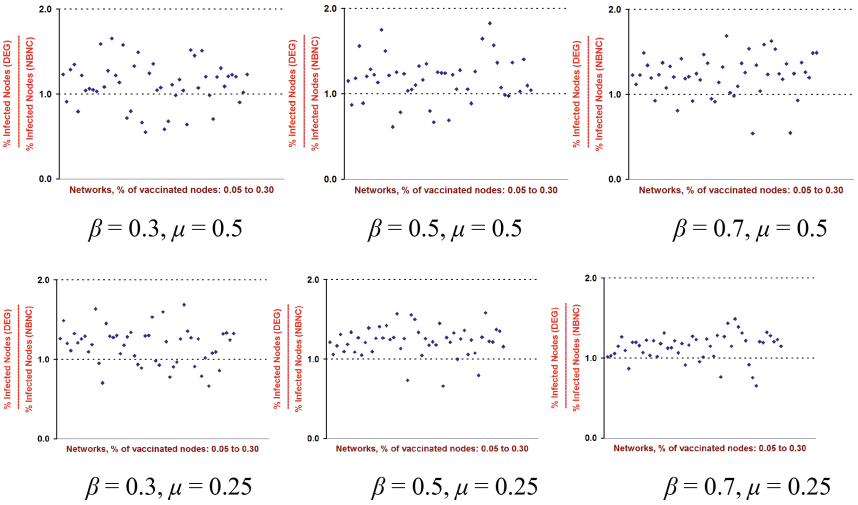


Fig. 4. Ratio of the % Infected Nodes for the 10 Real-World Networks and for different Values of the Fraction of Nodes that are Vaccinated with DEG and NBNC as Node Selection Strategies for Vaccination

5 Related Work

In a recent work [1], centrality metrics were considered for vaccination to contain the spread of an infection per the SIR model. Unlike our work, in [1], the nodes chosen for vaccination and the nodes that recover from the infection are simply removed from the network; such a simulation approach will not work for the SIS model as (with the SIS model) the recovering nodes immediately become susceptible and need to be considered part of the network as a potential candidate for future infections. In [2], the authors observed that betweenness centrality [3] is effective in containing the spread of an infection in synthetic networks generated per the Barabasi-Albert scale-free model [4]; whereas, the degree centrality was found to be more effective in reducing the spread of an infection in real-world networks. Note that betweenness centrality is a computationally-heavy metric and its computation needs to be synchronous as well as requires global knowledge. Recently [9, 10], community-aware approaches have been evaluated for preferential vaccination of nodes; but, such strategies require the knowledge of global information all the time. On the other hand, the NBNC tuple [11] for a node can be computed using local neighborhood information of the node and its neighbors. Non-centrality and/or non-community-based strategies for vaccinating nodes have been typically found to be less effective [12] (for example: the strategy [13] of choosing the random neighbor of a randomly chosen node).

Per [5], the problem of identifying critical nodes for vaccination and protection of susceptible non-vaccinated nodes has been shown to be equivalent to the problem of identifying the super spreader nodes for information diffusion [6]. In [7], the authors showed that given an adjacency matrix-style contact matrix of people in a social network (without any constraint on the structure of the matrix), the diffusion probability of the information has been observed to be proportional to the largest Eigenvalue (a.k.a. the spectral radius [8]) of the adjacency matrix; the lower the spectral radius, lower the diffusion probability of the information (also applicable to viral diffusion).

6 Conclusions and Future Work

We propose a simulation-based approach to identify the most effective centrality-based node selection strategy for preferential vaccination of nodes to reduce the fraction of infected nodes during any spread of an infection. We used the SIS model for the infection spread. The simulation and evaluation procedures employed in this paper can be used for any measure adopted for preferential vaccination of nodes. We observe the neighborhood-based bridge node centrality (NBNC) approach to be more effective compared to the widely considered degree centrality (DEG) approach for preferential vaccination. We thus show that the bridge node-based NBNC tuple is more effective in identifying node(s) whose neighbors are less likely to infect each other if the node(s) are vaccinated. This is the key factor behind the success of NBNC over DEG towards reducing the number of infected nodes per round, especially with increase in the fraction of vaccinated nodes. As part of future work, we plan to evaluate the effectiveness of NBNC vs. DEG for preferential node vaccination with respect to the familiar infection spread models such as the SIR and SIRS models [17] involving both real-world networks as well as synthetic networks.

Acknowledgement. The work leading to this paper was partly funded through the U.S. National Science Foundation (NSF) grant OAC-1835439 and partly supported through a sub contract received from University of Virginia titled: Global Pervasive Computational Epidemiology, with the National Science Foundation as the primary funding agency. The views and conclusions contained in this paper are those of the authors and do not represent the official policies, either expressed or implied, of the funding agency.

References

1. Sartori, F., et al.: A comparison of node vaccination strategies to Halt SIR epidemic spreading in real-world complex networks. Sci. Rep. **12**(21355), 1–13 (2022)
2. Wei, X., Zhao, J., Liu, S., Wang, Y.: Identifying influential spreaders in complex networks for disease spread and control. Sci. Rep. **12**(5550), 1–11 (2022)
3. Brandes, U.: A faster algorithm for betweenness centrality. J. Math. Soc. **25**(2), 163–177 (2001)
4. Barabasi, A.-L., Albert, R.: Emergence of scaling in random networks. Science **286**(5439), 509–512 (1999)
5. Paluch, R., Lu, X., Suchecki, K., Szymański, B.K., Hołyst, J.A.: Fast and accurate detection of spread source in large complex networks. Sci. Rep. **8**(1), 1–10 (2018)

6. Zhang, D., Wang, Y., Zhang, Z.: Identifying and quantifying potential super-spreaders in social networks. Sci. Rep. **9**(14811), 1–11 (2019)
7. Wang, Y., Chakrabarti, D., Wang, C., Faloutsos, C.: Epidemic spreading in real networks: an eigenvalue viewpoint. In: The 22nd International Symposium on Reliable Distributed Systems, pp. 25–34 (2003)
8. Guo, J.-M., Wang, Z.-W., Li, X.: Sharp upper bounds of the spectral radius of a graph. Discret. Math. **342**(9), 2559–2563 (2019)
9. Cherifi, H., Palla, G., Szymanski, B.K., Lu, X.: On community structure in complex networks: challenges and opportunities. Appl. Netw. Sci. **4**(117), 1–35 (2019)
10. Ghalmane, Z., Cherifi, C., Cherifi, H., El Hassouni, M.: Centrality in complex networks with overlapping community structure. Sci. Rep. **9**(10133), 1–29 (2019)
11. Meghanathan, N.: Neighborhood-based bridge node centrality tuple for complex network analysis. Appl. Netw. Sci. **6**(47), 1–36 (2021)
12. Lev, T., Shmueli, E.: State-based targeted vaccination. Appl. Netw. Sci. **6**(6), 1–16 (2021)
13. Gallos, L. K., Lilijeros, F., Argyrakis, P., Bunde, A., and Havlin, S.: Improving immunization strategies. Phys. Rev. E. **75**, 045104 (2007)
14. Newman, M.: Networks: An Introduction¸ Oxford University Press, Oxford, UK (2010)
15. Ma, J., van den Driessche, P., Willeboordse, F.H.: The importance of contact network topology for the success of vaccination strategies. J. Theor. Biol. **325**, 12–21 (2013)
16. Musiał, K., Juszczyszyn, K.: Properties of bridge nodes in social networks. In: Nguyen, N.T., Kowalczyk, R., Chen, S.M. (eds.) Computational Collective Intelligence. Semantic Web, Social Networks and Multiagent Systems. ICCCI 2009. LNCS, vol. 5796, pp. 357–364. Springer, Berlin (2009). https://doi.org/10.1007/978-3-642-04441-0_31
17. Liu, J., and Xia, S.: Computational Epidemiology: From Disease Transmission Modeling to Vaccination Decision Making, Springer (2020)
18. Fiedler, M.: Algebraic connectivity of graphs. Czechoslov. Math. J. **23**(98), 298–305 (1973)
19. Strang, G.: Linear Algebra and its Applications, Brooks Cole, Pacific Grove, CA, USA (2006)

Social Media Applications' Privacy Policies for Facilitating Digital Living

Kagiso Mphasane, Vusumuzi Malele(✉), and Temitope Mapayi

Unit for Data Science and Computing, School of Computer Science and Information Systems (CSIS), North-West University, Vanderbijlpark Campus, Vanderbiljpark, South Africa
vusi.malele@nwu.ac.za

Abstract. Mobile devices are widely utilised over the globe to connect people with their loves, social interactions, or business. The mobile devices are rapidly increasing, creating business opportunities for third-party applications developers. Unfortunately, the increase of third-party applications also creates an increase security and privacy issues. Mobile applications privacy policies ensure that one or users of the applications understand how their personal information is collected by the developer of that mobile app. Acceptable usage policies outline those actions which are not allowed or allowed when using the application. Some acceptable usage policies contain disclaimers of liability for the mobile application developer. This paper compares privacy policies of four popular mobile applications and assess their acceptable usage within the context of third-party applications on mobile users. The primary data was collected through questionnaires that gather the understanding and perception of people regarding third-party policy. The secondary data collection was used for comparing the third-party application's policies. The results show that most people disregard policies and they do not understand the policies. Furthermore, all the four third-party applications use both the API and cookies features. The privacy policies API feature in this case is used to allow the third-party companies to access various features like user data and posting functionality. With policy that these third-part companies are using, and the increase in the usage of smartphone third-party applications for mobile users' daily interactions with third-party applications; then there is a need for recommending ways of safeguarding the mobile device users.

Keywords: Privacy policies · cookies · API · third-party app learning

1 Introduction

In 2023, the number of iOS app users on the Apple App Store was 1.8 million, and the number of Android apps on the Google Play Store was 2.4 million, respectively. The increase in the number of users provides a platform for an increase in threats, vulnerabilities, and violations that could infringe the mobile users' privacy. Figure 1 below illustrate the extent in which the threats and vulnerabilities has increased in 2022 and the first quarter of 2023.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 37–48, 2024.
https://doi.org/10.1007/978-3-031-53552-9_4

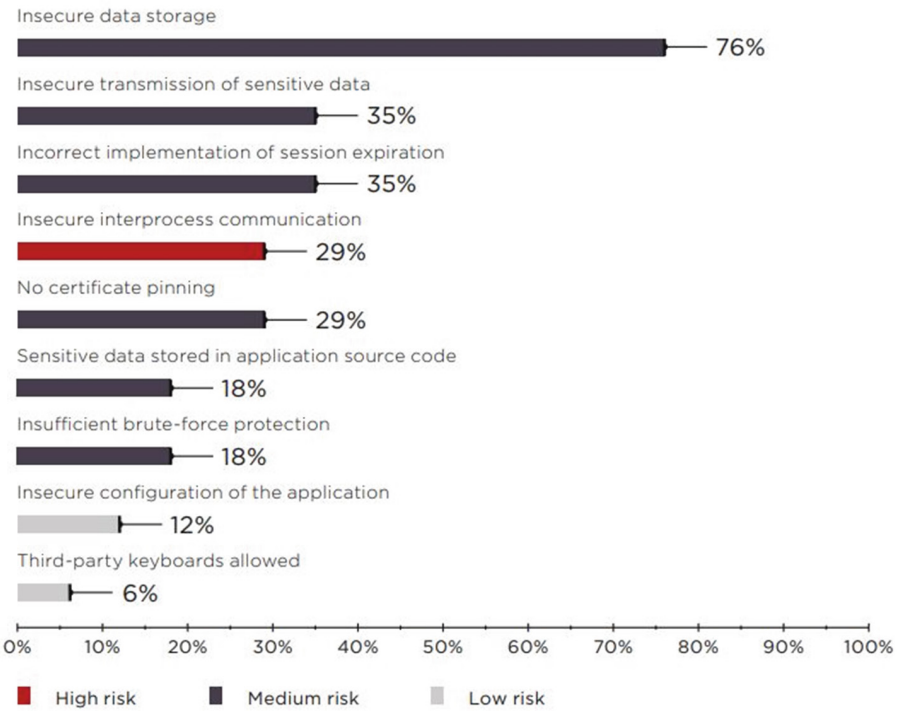


Fig. 1. Mobile application threats (Source: [1])

The increased in mobile device users has created a platform for third-party application (apps) to increase. The third-party app is a software app developed and designed to connect with another service to either provide enhanced features or access profile information. In this era, people are living in modern communities were almost everyone is profiled, using the data and information collected by third-party apps without users consent or awareness [2–4].

Given the extensive range of mobile apps accessibility, each with a different level of security and privacy, it is crucial to identify which third-party apps are used by individuals and can place the users at risk [4]. Furthermore, each user should be aware of the data that these apps are collecting [4]. In this regard, the digital age privacy policy for safeguarding the users is a vital research area.

Despite this Introduction, Sect. 2 describes some literature review. Section 3 describes the methods that were used to conduct this study, Sect. 4, presents the results and discussion. Then Sect. 5 presents the recommendation and conclusion.

2 Literature

These days personal and intimate details are provided to various companies with or without consent since most users do not concentrate on mobile apps policies. The mobile users become naïve thinking that the third-party organizations will keep their information

safe and secure [2]. Paul Bischoff, one of the privacy experts for Comparitech and a regular commentator on cyber security and privacy topics in national and international media including New York Times, BBC, Forbes, The Guardian, and many others. His expert opinion indicated that Apple Finder app has several flaws, one of them is to track user's location which go against privacy policies in most of the countries [5].

Some of the most popular mobile application in the world are TikTok, Tinder, YouTube, Twitter, Facebook, Instagram, and Spotify. However, four widely used third-party apps in South Africa are Tinder, Facebook, Instagram, and Spotify [2]. Often without people consent or awareness some of these apps mine the users' data. For example, Tinder has the ability of detecting individuals' movement, able to mine users' digital existence, and preferences [5].

A huge amount of personal and sensitive data is shared on Tinder, Facebook, Instagram, and Spotify making them number one target for attackers. When analysing user perception on Facebook; an analysis revealed significant mismatches between users' privacy perceptions and reality. For example, were too optimistic through their perceptions of information collection, but also on their self-efficacy in protecting their information [6].

In the case of Tinder, users are more concerned about institutional privacy than social privacy [7]. Mobile device users provided different motivations for using Tinder, as such these motivations affect social privacy concerns more strongly than institutional concerns. With respect to Tinder, loneliness significantly increases users' social and institutional privacy concerns.

Mobile device users have increased the usage of smartphone third-party applications for their daily interactions and other purposes [3, 4] such an increase, give third-party apps an opportunity for business and as well as possible user violation. Hence, their policies need to be investigated. In lieu of this, this paper compares privacy policies and the acceptable usage policies for four popular mobile third-party apps: Tinder, Facebook, Instagram, and Spotify.

3 Research Methodology

The research onion highlighted in Fig. 2 was used to guide this paper. Kaushik and Walsh [8] writes "*Pragmatism as a research paradigm finds its philosophical foundation in the historical contributions of the philosophy of pragmatism and, as such, embraces plurality of methods. As a research paradigm, pragmatism is based on the proposition that researchers should use the philosophical and/or methodological approach that works best for the research problem that is being investigated*". In this regard, this paper adopted a pragmatic research philosophy because the approach aim was to look at the policy issues in practical point of view.

This paper adopted a deductive research approach, since policy is a top-down approach in which third-party app organisation interact with the users before deploying their app to their mobile devices. Since, the mobile users need their respective app, third-party organisation provide them with a "take-or-leave" approach, making it a top-down approach.

A research design chosen by this study is the case study approach of four third-party apps. It investigates the privacy and security policy of four third-party apps, Tinder,

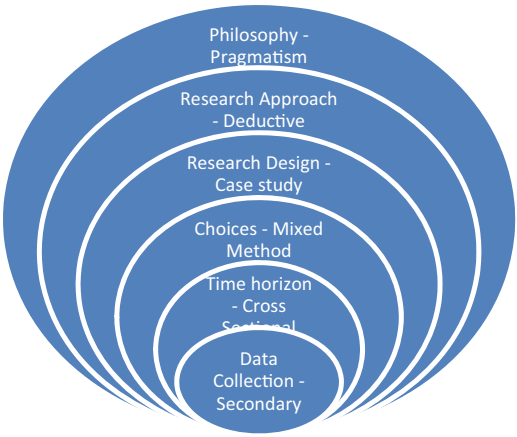


Fig. 2. The research onion adopted in this paper.

Facebook, Instagram, and Spotify. In this regard, the Policy analysis method which uses both qualitative and quantitative methods is used in this study. The latter is the mixed methods research in which third-party apps' policy analysis (i.e., secondary data) is a qualitative approach and the quantitative approach was done by collecting primary data from the group of 33 random mobile users. The distributed questionnaire collect data from November 2022 to February 2023. The survey aimed to obtain insights about participants smartphone usage patterns, and their attitudes about privacy concerns in context of mobile apps. The survey comprised of three sections, background data, technical knowledge, and privacy awareness. The questionnaire asked five questions:

- Do you always accept the policy of the apps without reading them?
- Are apps' policies easy to understand?
- Are apps' policies easily readable?
- Do you think free and open-source apps can be a convenient and inexpensive way to add extra functionality to your experience with mobile device?
- Do you think the third-party apps can infect your mobile device with malicious software?

Finally, to draw important conclusions, the facts must be presented properly. The author therefore looked at each participant's response and Software from IBM SPSS Statistics was used to analyze the survey data. The findings of the data is described below.

4 Findings

4.1 Understanding and Perception

The C.I.A triad is a framework that points to the users' threats and intrinsic complications which might cause data leakage and jeopardize the information's confidentiality, integrity, and availability. The latter talks to the necessary need for the protection of

users' sensible data and Personally Identifiable Information (PII) that could be accessed through the mobile app. In this regard, the author used the C.I.A triad as its theoretical benchmark for assessing the privacy and security issues that emanates from third-party apps policies and its violations. The data displayed below are responses to surveyed questions.

Table 1. Randomly selected participants.

Age	Car Wash		Church		Mall		Total
	M	F	M	F	M	F	
12 to 20	1	0	3	4	2	2	12
21 to 30	6	2	3	3	4	5	23
31 to 40	5	1	2	4	2	4	18
40+	2	2	3	3	1	1	12
Total	14	5	11	14	9	12	65

Five pilot study questions were asked from 65 randomly selected mobile users from three social areas: car wash, church, and shopping mall. Table 1 represent the number of participants, their demographics, and the area where data was collected. Of the 65 people, 19 were sampled from a car wash facility, 25 from church, and 21 from the shopping mall. Figures 3, 4, 5, 6, and 7 provided the responses to the pilot questions.

Figures 3, 4 and 5 illustrate the understanding of participants on third-party policies. While Figs. 6 and 7 tested the participants perceptions about the cost and security of third-party apps.

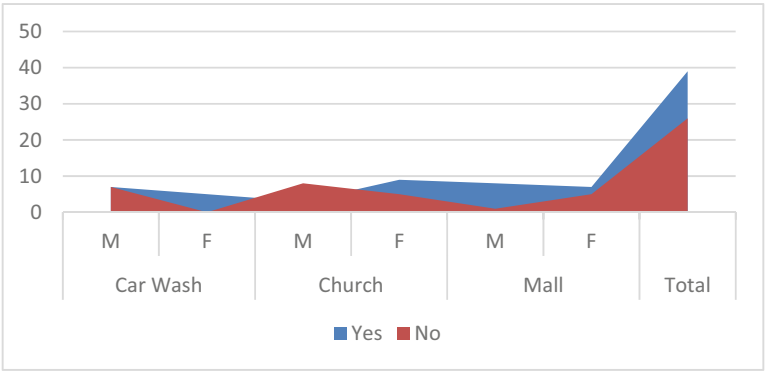


Fig. 3. Just accept the policy of the third-party apps without reading.

Figure 3 illustrates that majority of people always accept the policy of the third-party apps without reading them. This is corroborated by Figs. 4 and 5 that shows that

majority of participants thought that third-app's policies are not easy to understand and not readable, respectively.

Figures 4 and 5 that shows that majority of participants thought that third-app's policies are not easy to understand and not readable, respectively.

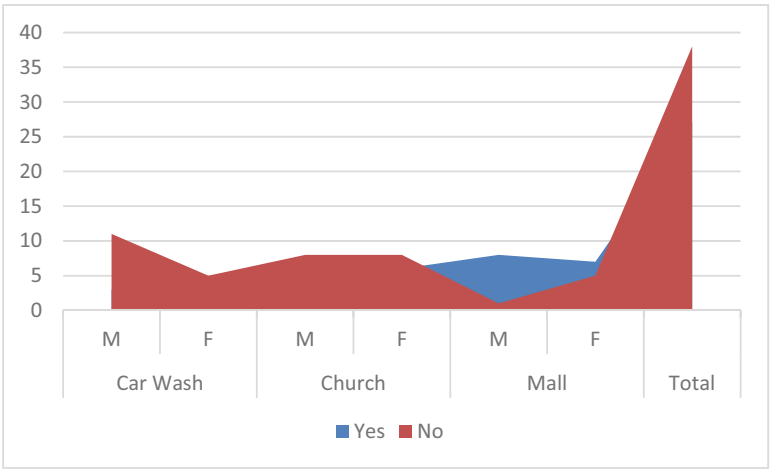


Fig. 4. Participants that thought third-party app's policies are easy to understand

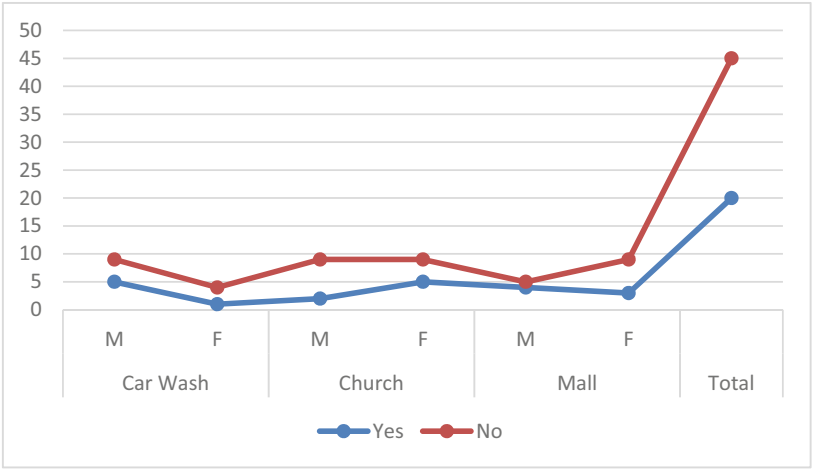


Fig. 5. Participants that thought third-party app's policies are readable.

Figure 6 showed that majority of participants feel that free and open-source apps can be a convenient and inexpensive way to add extra functionality but some of them are not adding value to their experiences. Figure 7, present a very interesting finding. It seems like the particpants were equaly likely to accept the fact that third-party apps could infect their mobile phones.

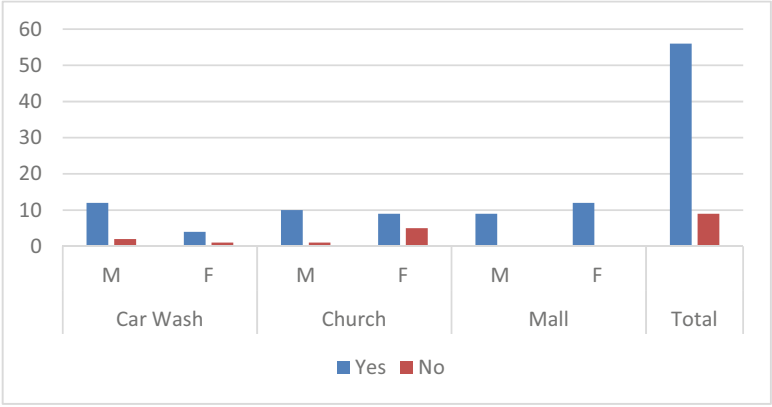


Fig. 6. Free and open-source apps can be a convenient and inexpensive.

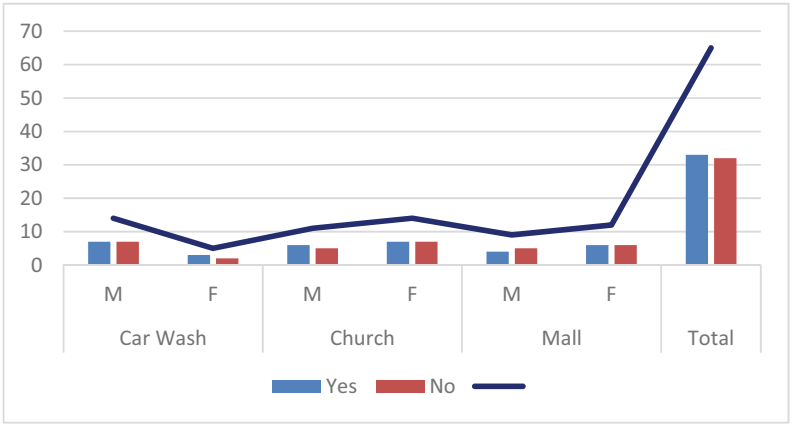


Fig. 7. Third-party apps can infect devices with malicious software.

Whatever analysis and observations are presented by Figs. 3, 4, 5, 6 and 7, their general contribution is that they show a need to unpack the third-party apps policy space, cost and security issues towards mobile users. For example, Fig. 7 illustrates a need for awareness of regarding the security matters of the third-party apps. Since this paper is an extract of a Master's dissertation, it is envisaged that the unpacking will be illustrated in the dissertations itself. The survey questions supplied to participants should generate replies that will allow the study goal to be met. Finally, proper presentation of the data analyzed is required to derive significant conclusions and provide recommendations based on the trends and patterns derived from the data collection.

5 Recommendations

To ensure the security of mobile apps, effective risk management for third-party applications entails assessing the overall potential risks associated with each third-party relationship, putting in place the appropriate controls and safeguards, and continuously keeping an eye out for potential threats and vulnerabilities. For consumers to ensure that no private information is accessed by unauthorized users, some tips include:

Review the data access request: All third-party apps should have defined security criteria that include monitoring and reporting obligations, data security and privacy standards, and incident response methods.

Update your mobile apps frequently: It is now common knowledge that third-party applications have developed a method for introducing patches for security vulnerabilities, ensuring that users cannot access other users' data through those vulnerabilities.

Evaluation of the developers' security practices: Users of mobile devices are advised to understand the security protocols and security measures before downloading any applications to make sure that no unauthorized users can access their personal data.

Implementation of role-based access control: It has been noted that third-party programs occasionally need user information to enhance their functionality and provide users with better services.

The report's suggestions will inform users of the numerous dangers they may encounter if vendors are given access to any sensitive data as well as how to secure communication on third-party applications.

5.1 Privacy Policies: Tinder

Tinder is one of the most popular mobile app. The Mobile application, which is owned and operated by the Match Group, a limited company located in the USA. The application is an online dating app which uses geosocial networking application. The app policy has been written in English but has been translated into other languages [9].

The privacy policy only applies to those who use the application and Tinder services only [10]. Additionally, the company has indicated that Tinder privacy policy forms part of organisation's terms of use. The Match Group company has indicated that by using Tinder mobile app it means that one must abide by the privacy practices described in Tinder policy. One of the major concerns with using the application, it means that the company can collect and retain your personal information which might contribute to identifying the user of the app.

Lastly, if a user of the application is in the European Union region, then they are covered under the General Data Protection Regulation rights (GDPR). These rights include the right to be informed about collection of their data and right of access to one personal data [10].

5.2 Privacy Policy: Facebook

Facebook is one of the social media mobile applications owned by Meta company. The company updated application's privacy policy to reflect the current new name which was unveiled on 4 January 2022. From application's privacy policy the mobile application

collects a wide range of data from the users of the app like phone number, name, and email address. Furthermore, the application collects user activity such as shares, comments, and likes [11].

The data collected is used by the company to provide personalised content to users. The data assist users to see what the services and products they like one Facebook apps. The Facebook's privacy policy user data can be shared with a third party without user consent to provide a better advertising. The latter means that the mobile app use cookies and similar technologies to collect their mobile app user activity, behavior, and user activity [11].

5.3 Instagram

Instagram mobile app is also one of the social media applications owned by Meta company. Just like Facebook, the company updated their privacy policy for the Instagram to reflect the new company in forth January 2022. On Instagram's privacy policy, meta has described what type of information that they collect from the mobile social app.

Different from Facebook, Instagram app collects information depending on the content that one provides. From the privacy policy, the type of information that app collects include one that one provides during sign-up and when creating or sharing content and messages to the other users. If the information in this case was the provision of the photo, the data collected will include even photo metadata like the location of the photo or when the photo was provided. Furthermore, the application collects information about people's networks and connections [12].

Just like the other three mobile apps, Instagram app also use tracking cookies to assist the company in improving and personalising content services and providing a safer experience. These cookies can also be used to remember changes to text size, font, and other parts of pages that you can customize. These technologies can remember when your device visited a website or service, and they may also be able to track your device's browsing activity on websites or services other than Instagram. This information may be shared with organisations other than Instagram, such as advertisers and/or advertising networks, to deliver advertising and help measure the effectiveness of a campaign.

5.4 Spotify

Different from other application, Spotify is a digital music app which gives users of the application access to millions of songs from songs creators all over the world. Spotify's privacy policy mainly specifies how they process the users of application at the company. According to Spotify, the policy applies to the use of all company streaming services as a client or the user of the application. It applies to Spotify services such as customer support, Spotify websites and community Site [13].

Different from other privacy policies described previously, Spotify privacy is not Spotify terms of use. From the company's privacy policy, the mobile application offers users right access to their personal information; this includes even those who are located outside EU boundaries. Some of the information that the mobile application collects about their users as indicated in their privacy policy include user data that is during sign-up.

From the company privacy policy, the type of data collected from users depends on the type of service that users subscribe to. The mobile application also collects street address information. The mobile application asks for users' street address for tax administration and so that they can deliver gifts or physical goods that one requests. Lastly, just like other mobile apps the company employs the use of cookies for content personalization [13].

5.5 Similarities

All the four mobile apps; Tinder, Facebook, Instagram, and Spotify collect almost similar type of information such as users' names, their email address, and their phone numbers. Second, from the description of their privacy policies all the four mobile apps use both API and cookies features. API feature in this case is used to allow the mobile apps companies to access various features like user data and posting functionality.

From Instagram privacy policy they have described that they have integrated their application with Facebook app. This means that the two apps use APIs feature to allow that interaction between the two apps.

Similarly, Spotify uses APIs to allow access to manage application's playlists and to incorporate the various functionalities. Other apps not included in this study such as Finder uses APIs so that they can access information from other sources like banks. This would allow the application to provide information in real-time without having to build information from scratch.

As highlighted all the mobile apps uses cookies feature to enable creation of personalized content for their users. Table 2 show the similarities on what how the different apps collect data from users (Table 3).

Table 2. Similarities of the Four Third-Party App.

Data Collection	Tinder	Facebook	Instagram	Spotify
Collect Username during account creation	✓	✓	✓	✓
Collect Email during account creation	✓	✓	✓	✓
Collect Password during account creation	✓	✓	✓	✓
Collect Physical address				✓
Device Identifiers	✓	✓	✓	✓
Use of API	✓	✓	✓	✓
Log File Information	✓	✓	✓	✓
Use of Tracking cookies	✓	✓	✓	✓
Sharing of Data to Third Party Agencies	✓	✓	✓	✓

By using these app services, you acknowledge and accept that you are provided with a platform to publish content, including images, comments, and other materials ("user content"), to the service and to publicly distribute user content. This implies that any of

Table 3. Acceptable usage policy.

Acceptable usage policy	Tinder	Facebook	Instagram	Spotify
Contents of their usage policy	Does not allow users to engage in any illegal actions this includes money laundering or money fraud. Also, users are required not to attempt to disrupt app services in form of hacking or in any other way	Does not allow uses of the app to engage in any form of bullying or actions which violates application's community standards	Does not allow users of their app to engage in any form of harassing or hate speech	Does not allow one to share their account to others Does allow engagement that violates their terms of service

your user content that you make publicly available through the service may be searched for, seen, used, or shared by other users in accordance with the terms and conditions of this privacy policy and terms of use provided by these apps. The table below summarizes these apps' acceptable usage policies.

6 Conclusion

First, as one can note from the described privacy policies for the four mobile apps; Tinder, Facebook, Instagram, and Spotify they collect almost similar type of information such as users' names, their email address, and their phone numbers. Second, from the description of their privacy policies all the four mobile apps use both API and cookies features. The API feature in this case is used to allow the Mobile apps companies to access various features like user data and posting functionality. According to Instagram's privacy policies, they have integrated their app with the Facebook app. This means that the two apps use APIs feature to allow interaction between the two apps. Similarly, Spotify uses APIs to allow access to manage its playlists and to incorporate the various functionalities. Also, Tinder makes use of APIs to get data from other sources on the mobile devices of other user profiles. This would allow the application to provide information in real-time without having to build information from scratch; and, as highlighted, all the mobile apps use cookies feature to enable creation of personalised content for their users.

Nevertheless, differences in their privacy policies comes from acceptable usage policies. Example for Spotify acceptable usage policy does not allow one to share their account to others or to use their application to engage in behaviors which violates their terms of service. On the other Instagram mobile application acceptable usage policy does allow users of their app to engage in any form of harassing or hate speech while Facebook app acceptable usage policy does not allow its uses of the app to engage in

any form of bullying or actions which violates its community standards. Lastly, Tinder app acceptable usage policy does allow its users to engage in any illegal actions this includes money laundering or money fraud. Also, users are required not to attempt to disrupt app services in form of hacking or in any other way.

References

1. Schneider, M., Chowdhury, M.M., Latif, S.: Mobile devices vulnerabilities. EPiC Ser. Comput. **82**, 92–101 (2022)
2. Goldstein, K., Tov, S.O., Prazeres, D.: The Right to Privacy in the Digital Age. Pirate Parties International Press, Brussels (2018)
3. Bakopoulou, E., Shuba, A., Markopoulou, A.: Exposures exposed a measurement and user study to assess mobile data privacy in context (2020). arXiv preprint arXiv: 2008.08973
4. Hayes, D., Cappa, F., Le-Khac, N.A.: An effective approach to mobile device management: security and privacy issues associated with mobile applications. Digit. Bus. **1**(1), 100001 (2020)
5. Brodsky, S.: Apple's device finder app could expose you, experts say (2021). https://www.lifewire.com/apples-device-finder-app-could-expose-you-experts-say-5116026. Accessed 29 March 2023
6. Seng, S., Al-Ameen, M.N., Wright, M.: A look into user privacy andthird-party applications in Facebook. Inf. Comput. Secur. **29**(2), 283–313 (2021). https://doi.org/10.1108/ICS-08-2019-0108
7. Lutz, C., Ranzini, G.: Where dating meets data: investigating (2017)
8. Social and institutional privacy concerns on tinder. Social media + Society, pp. 1–12. https://doi.org/10.1177/2056305117697735
9. Kaushik, V., Christine, A.: Walsh pragmatism as a research paradigm and its implications for social work research. Soc. Sci. **8**(9), 255 (2019). https://doi.org/10.3390/socsci8090255
10. McDonald, A.M., Reeder, R.W., Kelley, P.G., Cranor, L.F.: A comparative study of online privacy policies and formats. In: Goldberg, I., Atallah, M.J. (eds.) Privacy Enhancing Technologies. PETS 2009. LNCS, vol. 5672, pp. 37–55. Springer, Berlin (2009). https://doi.org/10.1007/978-3-642-03168-7_3
11. Match Group, 2022. Privacy and Cookies Policy. https://mtch.com/privacy. Accessed 30 March 2023
12. Johnson, M.S., Egelman, S., Bellovin, S.M.: Facebook and privacy: it's complicated. Proc. Eighth Symp. Usable Priv. Secur. **1**(2), 1–15 (2022). https://doi.org/10.1145/2335356.2335369
13. Instagram, 2022. Instagram help center: data policy. https://help.instagram.com/155833707900388. Accessed 30 March 2023
14. Spotify (2021). Spotify privacy policy. https://www.spotify.com/kr-en/legal/privacy-policy/. Accessed 30 March 2023

Correlation Analysis of Student's Competencies and Employ-Ability Tracer Study of Telkom University Graduates

P. H. Gunawan[1(✉)], I. Palupi[1], Indwiarti[1], A. A. Rohmawati[1], and A. T. Hanuranto[2]

[1] Human Centric (HUMIC) Engineering, School of Computing, Telkom University, Bandung, Indonesia
{phgunawan,irmapalupi,indwiarti,aniqatiqi,athanuranto}@telkomuniversity.ac.id

[2] School of Electrical Engineering, Telkom University, Jl. Telekomunikasi No 1, Terusan Buah Batu, Bandung 40257, Indonesia

Abstract. A tracer study is an essential tool for universities to analyze education implementation as indicated by alumnus performance in the industry. This research elaborates on correlation analysis between students' competencies obtained during study at Telkom University and waiting time for the first job. This correlation is essential to investigate how the suitability of education accelerates alums' careers. Our results show that suitability of education and time waiting are highly correlated, with more than 90%. Here, the competency's importance is ranked. It shows that course knowledge, hard skills, and communication are the three most essential competencies in Telkom University alums' professional careers. Moreover, to validate the consistency of alumni perception used in data analysis with industry opinion as alumnus employer, a consistency score of 93.63% was computed, which means they are consistent. Furthermore, in this research, the logit model used in analyzing the survey has several key benefits. It allows the researchers to examine the intricate connections between different variables and outcomes, providing valuable insights into the factors that shape graduates' career paths and achievements.

Keywords: Correlation · tracer study · Spearman's rank-order · logit regression

1 Introduction

Tracer study is a powerful tool to evaluate the performance of study programs, faculty, and universities from graduated students [6,9]. Its abilities involve tracking the employment status of graduates and gathering information on their skills, knowledge, and experience, which can be used to improve the curriculum and prepare future graduates for the job market. In the tracer study, some questions

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 49–60, 2024.
https://doi.org/10.1007/978-3-031-53552-9_5

regarding waiting time, job suitability, first salary range, and employ-ability performance are delivered for graduate students. This information is helpful as an evaluation tool to examine the relevance between colleges and the world of business and industry. According to the Ministry of Education in Indonesia [4], the tracer study is aimed at tracing the traces of alums which is carried out two years after graduation and aims to find out:

(i) The educational results encompass the transition from higher education to the workforce.
(ii) The educational output involves self-evaluation of skill mastery and competence acquisition.
(iii) The educational procedure entails assessing the learning process and the contribution of higher education to competence acquisition.

The outcomes of the tracer study provide universities with information about the employment status of their graduates and enable them to tailor their educational programs to meet the specific competencies required in the workforce. By reporting the tracer study results to Higher Education, it aids the Government in aligning their programs with the demands of the job market, thus effectively mapping the educational development in Indonesia to meet the needs of the working world.

Telkom University, under the Directorate of Career, alums, and Endowment (CAE), has been to collect some information from alums by using a tracer study in 2022. More valuable information may not directly be obtained from a sample of individuals survey of tracer study, so it requires methods or tools to extract the data and gain important information for future evaluation. The research question is essential in defining the survey material in tracer studies because it guides the entire research process and determines the information that needs to be collected. These research questions can be used to gather data on the employment and educational outcomes of graduates from the university. The information gathered can be used to improve the university's curriculum and programs and provide valuable information to prospective students and employers. By tracking the outcomes of its graduates, the university can ensure that it provides a quality education that prepares its students for success in their careers. The research questions such as salary rate, job fields, and career competitiveness are commonly included in tracer study question surveys. One of many ways to measure the employ-ability of graduates is based on their waiting time until they find a first job. In [1,2,7,8,12] they investigated the relevance of the adapted curriculum and the suitability of competencies they learned in college to the job.

This research aims to investigate which of the competencies gained by students from Telkom University education affect their career start using statistical analysis. We also examine the consistency score between the alumni perceptions and graduate users' perceptions of competencies that affect employ-ability in order to avoid self-reported alumnus opinions. The length of time for alums to land their first job is chosen as one of the employ-ability representatives. Therefore, statistical methods such as correlation metrics and regression are appropriate in this research. The paper is structured as follows. The methodology of this

research is explained in Sect. 2. The results and discussion are elaborated on in Sect. 3. Moreover, the conclusion of this research is given in Sect. 4.

2 Methodology

2.1 Spearman's Rank-Order Correlation

In this research, the dataset from tracer study survey is in categorical form. Thus, to show the correlation between two categorical data, Spearman's Rank-Order Correlation is used in this research. The Spearman's Rank-Order Correlation is given by [5,10,11]

$$\rho = \frac{(n^3 - n) - \sum_{i=1}^{n} d_i^2}{n^2(n-1)} \tag{1}$$

where d_i is difference in paired rank at i-th data, n is the total number of data. The value of (1) is between interval $[-1, 1]$. If $\rho \to 1$, then there is a positive associated rank. Meanwhile, when $\rho \to -1$ it means a negative associated rank and when $\rho = 0$, it means no associated rank.

2.2 Chi-Square Test of Independence

The chi-square test of independence is used to determine whether two categorical variables have a significant relationship. It is frequently used in social science and business research to analyze survey data, but it can be applied to any situation involving two categorical variables. The test compares the observed frequencies of each category combination to the frequencies that would be expected if the variables were independent. The test statistic is computed as follows:

$$\chi^2 = \sum \frac{O - E}{E}$$

where O represents the observed frequency and E represents the expected frequency for each category combination. The aggregate is calculated across all categories. Calculating the degrees of freedom for the chi-square test of independence:

$$df = (r - 1) \cdot (c - 1)$$

where r is the number of entries in the table and c is the number of columns.

The null hypothesis of the chi-square test for independence is that the two variables are unrelated. The alternative hypothesis is that the two variables are significantly associated. Construct a contingency table containing the observed frequencies for each category combination to conduct the test. Then calculate the expected frequencies for each category combination under the independence assumption. Finally, calculate the test statistic using the chi-square formula and compare it to the critical value derived from the chi-square distribution with the appropriate degrees of freedom. If the calculated chi-square value exceeds the critical value, reject the null hypothesis and conclude that the two variables

are significantly associated. If the calculated chi-square value is less than the critical value, the null hypothesis cannot be rejected, and there is no significant association between the two variables.

Notably, the chi-square test of independence is a non-parametric test that makes no assumptions about the data distribution. However, it does presume that the observations are independent and that the expected frequencies for each cell in the contingency table are greater than or equal to five. If the expected frequencies for some cells are less than 5, an alternative test, such as Fisher's exact test, may be necessary.

2.3 Logit Regression Correlation Analysis

When the dependent variable is nominally scaled, logistic regression is used as a subset of regression analysis. Logistic regression analysis is thus the scaled version of linear regression, in which the dependent variable of the regression model must be interval-scaled at the very least. It is possible to use logistic regression to explain the dependent variable or to estimate the probability of occurrence of the variable's categories. The relationship between dependent and independent variables in logistic regression is not linear. Consequently, the regression coefficients cannot be interpreted in the same way as in linear regression. An independent variable is considered good in linear regression if it strongly correlates with the dependent variable. In contrast, it is considered good in logistic regression if it allows the groups of the dependent variable to be distinguished significantly.

At the multinomial logistic regression, the categorical dependent variable Y has more than two possible outcomes, namely $\{1, 2, \cdots, K\}$.

Since the odds of each class $k = 1, 2, \cdots, K-1$ against class K following

$$\log \frac{P(Y=k)}{P(Y=K)} = \beta_{0k} + \beta_{1k}X_1 + \beta_{2k}X_2 + \cdots + \beta_{pk}X_p \tag{2}$$

the probability for each class $k = 1, 2, \cdots, K-1$:

$$P(Y=k) = \frac{e^{\beta_{0k}+\beta_{1k}X_1+\beta_{2k}X_2+\cdots+\beta_{pk}X_p}}{1+\sum_{k=1}^{K-1} e^{\beta_{0k}+\beta_{1k}X_1+\beta_{2k}X_2+\cdots+\beta_{pk}X_p}} \tag{3}$$

and for class K,

$$P(Y=K) = \frac{1}{1+\sum_{k=1}^{K-1} e^{\beta_{0k}+\beta_{1k}X_1+\beta_{2k}X_2+\cdots+\beta_{pk}X_p}} \tag{4}$$

The estimation of $\boldsymbol{\beta}_k \in \mathbb{R}^{p+1}$ from a sample $\{(\mathbf{X_i}, y_i)\}_{i=1}^{N}$ is done by Maximum Likelihood Estimation (MLE). In the logistic model, the log-likelihood of $\boldsymbol{\beta}$ is

$$l(\boldsymbol{\beta}) := \sum_{i=1}^{N}\sum_{k=1}^{K-1}\left(y_{ij}(\beta_{0k} + \sum_{j=1}^{p}\beta_{jk}X_j)\right) - n_i \log\left(1 + \sum_{k=1}^{K-1} e^{\beta_{0k}+\sum_{j=1}^{p}\beta_{jk}X_j}\right) \tag{5}$$

and the ML estimation of $\boldsymbol{\beta}$ is

$$\hat{\boldsymbol{\beta}} = \arg \max_{\boldsymbol{\beta}_k \in \mathbb{R}^{p+1}} l(\boldsymbol{\beta}) \in \mathbb{R}^{K \times (p+1)}$$

2.4 Data Preprocessing

The survey data set consists of several variables which are obtained from the questions provided by CEA. There are 112 questions formed in several data types. For instance, the job suitability variable is qualitative discrete, and the waiting time variable is quantitative continuous. The distribution of the waiting period of alums to get their first job can be seen in Fig. 1. As we can see, its distribution is right-skewed and has some high outliers.

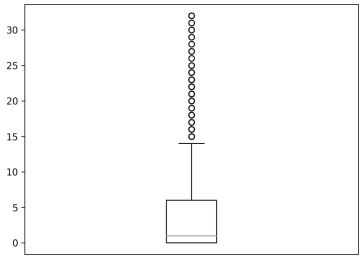


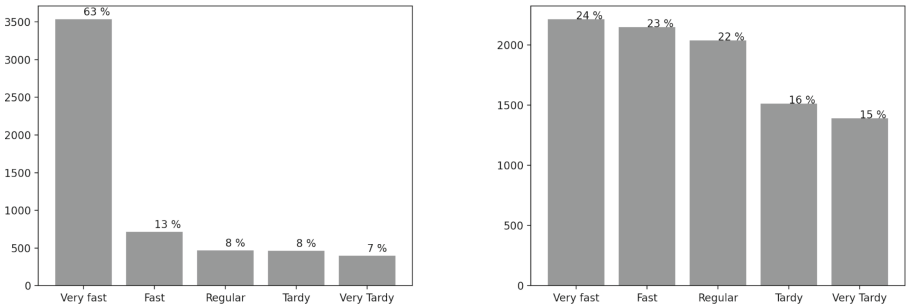
Fig. 1. Distribution of waiting time for the first job.

In order to compute the Spearman's order rank (1) of two variables, the type of the two variables should be in the same numerical format. The waiting time is then classified into five intervals shown in Table 1. The higher the class level, the faster a student gets a job. Also, Table 1 explains the categorical order for competency suitability related to the current job. The highest order means that alums consider their current job strongly suitable with their competencies background. Contrary, the Very weak suitable class means alums believe that their background is not supporting their current job.

Table 1. The level of job suitability and waiting time (t).

Ordering	Competency suitability	Time category
5	Strong Suitable	Very fast ($0 \leq t < 3$)
4	Suitable	Fast ($3 \leq t < 6$)
3	Moderate suitable	Regular ($6 \leq t < 9$)
2	Weak suitable	Tardy ($9 \leq t < 12$)
1	Very weak suitable	Very tardy ($t \geq 12$)

After classifying data in Fig. 1, the frequency distribution of waiting time can be found in Fig. 2(a). We found that around 63% of alums can find jobs quickly. From this number, we can observe that Telkom University students are very competitive in entering the career world. Since the portion of class 'very fast' is much more dominant than other classes, a resampling technique is required for the analysis model dealing with the imbalanced datasets.



(a) Before resampling.

(b) After resampling.

Fig. 2. Class distribution of dependent variable.

Imbalanced datasets conditions, as in Fig. 2(a), can lead to wrong predictions and cause a high bias in the resulting logit model. That is why a resampling technique is required for the component classes of the dataset to have a balance portion. After several experimental implementations of resampling methods, method SMOTEE-ENN [3] yields the best prediction accuracy. Figure 2(b) shows a decent balance dataset after resampling.

3 Results and Discussion

From Spearman's correlation coefficient in Fig. 3, all competencies have a strong positive correlation to the waiting time for getting the first job. Thus, alums with good competencies can get jobs faster than alums with fewer competencies.

Moreover, from Fig. 3, the waiting time has the highest correlation, which is 99%, with self-development competency, and contrary, the correlation between waiting time and argumentation competency yields the lowest correlation, about 86%. All competencies intensely matter to the waiting for alums to get a first job, based on the correlation coefficient value at Fig. 3. Next, a statistical significance test is applied to obtain additional information about the likelihood of the observed correlation occurring by chance and whether it is statistically significant. This test is used to determine whether the observed correlation is likely to have occurred by chance or whether it is statistically significant. Suppose that

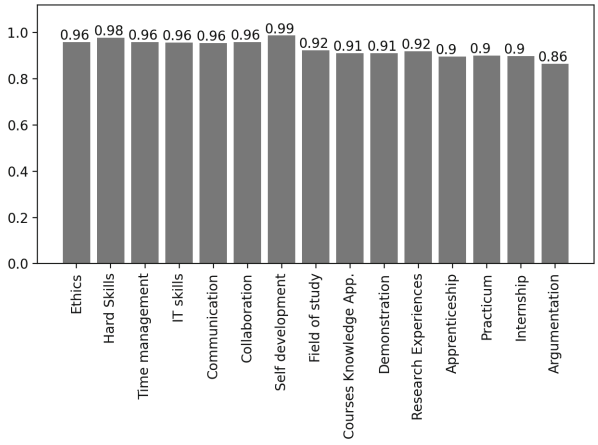


Fig. 3. Spearman's rank correlation of competencies to the waiting time.

the competencies, C_i for $i = 1, 2, \cdots, 15$, our hypothesis for significant test as follow:

H_0: C_i is not related to waiting time in the population
H_1: C_i is related to waiting time in the population.

The test's result can be seen in Table 2. All competencies have a p-value of less than 5% by using significant level $\alpha = 5\%$. Thus, it can be concluded that all competencies significantly affect the waiting time to get a job. However, the p-value of the field of study competency is much higher than the rest of the competencies, which is 1%, which means that despite using the significance level of 5%, this competency considerably relates to waiting time, it is not as significant as the others.

Compared to logit regression, Spearman rank-order correlation and the test of independence have drawbacks. They are limited in terms of the sorts of relationships they can study, their application to categorical outcomes, their inability to assess impact or account for confounding variables, their lack of probability estimation, and their scope for multivariate analysis. Logit regression provides a more flexible and complete technique for exploring complex relationships, and it is frequently used for evaluating this type of data.

Table 3 shows the logit regression report of the duration criteria for getting the first job after graduation, explained by the considered competencies. We can compare that regular and Very fast waiting time involves more competencies than Fast waiting time. Both have a significant relation with hard skills, internship, field of study, self-development, and apprenticeship. The regular waiting time regression coefficient has negative values regarding time management self-deviation, field of study, research experience, and internship. Moreover, the coefficient of a Very fast class of the waiting time also has a negative coefficient for

Table 2. The χ^2 test of independence for each competency to the waiting time, $df = 4$ and $N = 5580$ are all identic.

Competency	Report
Ethics	p = 0.00512
Hard Skills	p = 0.0
Time management	p = 0.0
IT skills	p = 0.0012
Communication	p = 0.00024
Collaboration	p = 0.0061
Self development	p = 0.00017
Field of study	p = 0.01306
Courses Knowledge App	p = 0.0
Demonstration	p = 9e-05
Research Experiences	p = 0.0018
Apprenticeship	p = 0.0
Practicum	p = 0.00102
Internship	p = 0.00016
Argumentation	p = 0.0

other variables such as IT skills, course knowledge, and argumentation. However, several competencies, such as hard skills, demonstration, and apprenticeship, reveal a positive coefficient on the regression model. Also, the competency field of study is noticed to have different directions; in terms of regular waiting time has a negative coefficient, the Very fast waiting time is vice versa. Furthermore, ethics only appear and positively correlate with the Fast waiting time, and IT skills only have a negative relation to the very fast waiting time.

Next, features importance is determined to measure the relative value or contribution of each predictor variable (feature) in predicting the result variable. This value assists in determining which features have the most significant influence on the model's predictions or results. In a logistic regression model, feature importance is often estimated by assessing the magnitude and direction of the coefficients associated with each predictor variable. These coefficients represent the estimated effect of the predictors on the log-odds of the outcome variable. Figure 4 is the feature importance order ranking of competency that influences the pace with which alums find work. The value is computed with only the weighted mean of the odds of the target class Regular, Fast, and Very fast (period of waiting 0–9 months).

The formula of importance of i-th competency is given in equation (6).

$$\text{importance}_i = 0.5\, e^{\beta_{i,\text{very fast}}} + 0.3\, e^{\beta_{i,\text{fast}}} + 0.2\, e^{\beta_{i,\text{regular}}} \tag{6}$$

Table 3. Multinomial logistic regression applied to all competencies.

y=Regular	Coef	Std err	z	p-value	95% Conf. Interval
Hard Skills	0.2233	0.067	3.317	0.001	[0.091, 0.355]
Time manage.	–0.1577	0.073	–2.174	0.030	[–0.300, –0.016]
Communication	0.3465	0.077	4.527	0.000	[0.196, 0.496]
Self dev.	–0.1533	0.078	–1.964	0.050	[–0.306, –0.000]
Field of study	–0.1966	0.076	–2.602	0.009	[–0.345, –0.049]
Demonstration	0.1901	0.067	2.821	0.005	[0.058, 0.322]
Research Exp.	–0.2364	0.064	–3.670	0.000	[–0.363, –0.110]
Apprenticeship	0.1577	0.062	2.525	0.012	[0.035, 0.280]
Internship	–0.1421	0.065	–2.186	0.029	[–0.269, –0.015]
y=Fast	Coef	Std err	z	p-value	95% Conf. Interval
Ethics	0.2331	0.075	3.104	0.002	[0.086, 0.380]
Time manage.	–0.3491	0.077	–4.530	0.000	[–0.500, –0.198]
Courses Knowledge	–0.1519	0.048	–3.197	0.001	[–0.245, –0.059]
Demonstration	0.3148	0.073	4.300	0.000	[0.171, 0.458]
Argumentation	–0.1651	0.071	–2.334	0.020	[–0.304, –0.026]
y=Very Fast	Coef	Std err	z	p-value	95% Conf. Interval
Hard Skills	0.5493	0.079	6.953	0.000	[0.394, 0.704]
IT skills	-0.1736	0.078	–2.230	0.026	[–0.326, –0.021]
Self dev.	–0.3425	0.088	–3.883	0.000	[–0.515, -0.170]
Field of study	0.1919	0.087	2.203	0.028	[0.021, 0.363]
Courses Knowledge	-0.1663	0.048	–3.458	0.001	[–0.261,–0.072]
Demonstration	0.1847	0.077	2.392	0.017	[0.033, 0.336]
Apprenticeship	0.3510	0.073	4.794	0.000	[0.208, 0.495]
Internship	–0.1514	0.075	–2.020	0.043	[–0.298, –0.004]
Argumentation	–0.6724	0.073	–9.222	0.000	[–0.815, –0.529]

From Fig. 4, it can be seen that the top five Competencies that considerably impact the immediacy of a student getting a first job after graduating from Telkom University are Hard skills, Course Knowledge, Research Experience, Ethics, and Communication. In the analysis of tracer study alumni, an investigation into what competencies have a dominant influence on employ-ability will be more valid if explored through a survey of graduate users. However, the sample size obtained from industry surveys is usually much smaller than the data collected from alumni. In this research, data from a number of alumni surveys was used so that it could adequately represent the condition of the graduate population. However, in order to avoid the possibility of self-reporting on the Tracer study data studied, validation is needed between alumni perceptions and graduate users' perceptions of competencies that affect employ-ability. The easiest

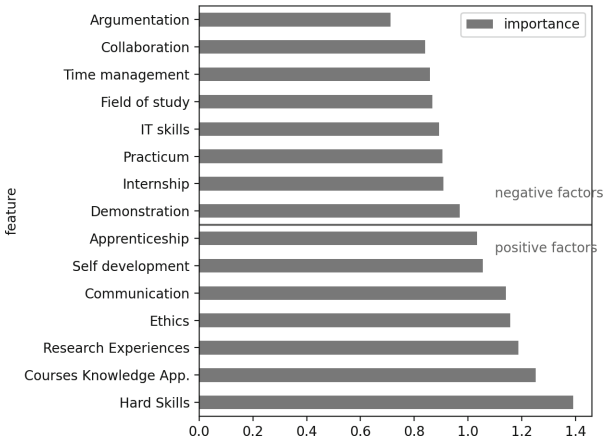


Fig. 4. Competency ranking that affects the duration for alumni to land a first job.

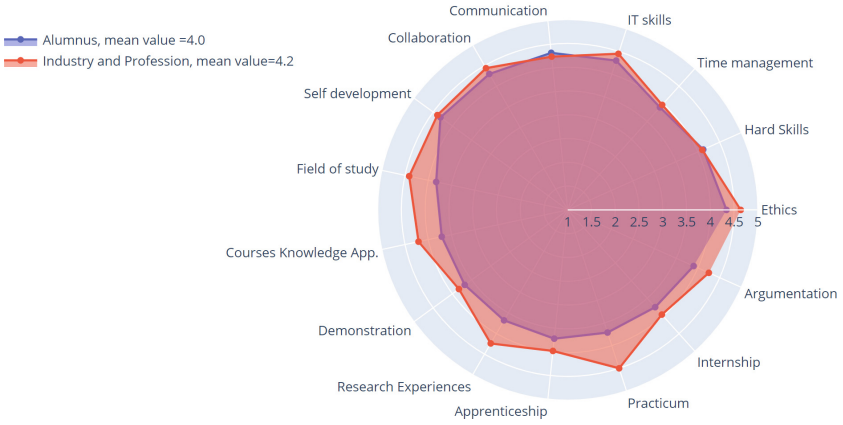


Fig. 5. Industry and Profession match of competency requirements.

way to assess is by the value of the correlation coefficient of the two perceptions for each competency. But again, due to the different sample sizes and to avoid using under-sampling techniques for the alumni survey, we propose a comparison of the mean values to assess the suitability of the perceptions of the two surveys.

On Fig. 5 we can see the mean comparison for each considered competency, with the value of mean and the standard deviation given in Table 4. The industry perception is important as the factual assessment of the alumnus's competency mastery level that we can see on Fig. 5 the values on average are almost entirely higher than alumni perceptions. To measure the suitability score of both perceptions, the formula (7) is used to compute the percentage consistency between them. We obtain the percentage consistency to be 93.63% of both perceptions,

which is high. Therefore, we can conclude that the assessment of the two is quite consistent so that the result in the Fig. 4 is valid with a score consistency of 93.63%, representing the assessment of the industry as the graduate user.

$$\text{Percentage consistency} = \left(1 - \frac{|\mu_{Industry} - \mu_{Alumnus}|}{\mu_{Industry}}\right) \times 100\% \quad (7)$$

Table 4. Mean and standard deviation of industry and alumnus perception from tracer study in 2022.

	μ	σ	Data size
Industry	4.24	0.58	860
Alumnus	3.97	0.85	5580

4 Conclusion

In summary, this article reports how to utilize Telkom University tracer studies in order to gain insight into the relevance between the taught curriculum and the world of business and industry. For more specifics from the survey of tracer studies, it was observed using correlation and logit regression analysis that the competencies of all provided competencies during the study were significantly overriding for the Telkom University alumnus to compete in the career world. The logit model's capability to handle categorical outcomes is particularly advantageous for tracer studies. In this case, we focus on determining the likelihood of a specific event, the time for a student to land their first job, with the competencies obtained during the study as the explanatory variables. The logit model provides a straightforward interpretation of results through odds ratios. These ratios quantify the influence of predictors on the likelihood of a particular outcome. Moreover, the logit model facilitates a clear understanding of the relative importance of different variables in shaping graduates' career outcomes. Policymakers, university administrators, and career counselors can leverage this information to make informed decisions and develop effective strategies for enhancing graduates' employ-ability.

References

1. Adriani, Z.A., Palupi, I.: Prediction of university student performance based on tracer study dataset using artificial neural network. Jurnal Komtika (Komputasi dan Informatika) **5**(2), 72–82 (2021)
2. Albina, A.C., Sumagaysay, L.P.: Employability tracer study of information technology education graduates from a state university in the Philippines. Soc. Sci. Hum. Open **2**(1), 100055 (2020)

3. Batista, G.E., Prati, R.C., Monard, M.C.: A study of the behavior of several methods for balancing machine learning training data. ACM SIGKDD Explor. Newsl **6**(1), 20–29 (2004)
4. Dikti: Tracerstudy Kemendikbudristek. https://tracerstudy.kemdikbud.go.id/ (2023). Accessed 23 February 2023
5. Gilmore, C., Barr, G., Dong, W.: Spearman's rank order coefficient. International Tables for Crystallography H (2019)
6. Gines, A.C.: Tracer study of PNU graduates. Am. Int. J. Contemp. Res. **4**(3), 81–98 (2014)
7. Heriyadi, B., et al.: Tracer study analysis for the reconstruction of the mining vocational curriculum in the era of industrial revolution 4.0. Turkish J. Comput. Math. Educ. (TURCOMAT) **12**(3), 3013–3019 (2021)
8. Jaeger, T.F.: categorical data analysis: away from ANOVAs (transformation or not) and towards logit mixed models. J. Mem. Lang. **59**(4), 434–446 (2008)
9. Kalaw, M.T.B.: Tracer study of bachelor of science in mathematics. Int. J. Eval. Res. Educ. **8**(3), 537–548 (2019)
10. Mohd Yani, A.A., Ahmad, M.S., Ngah, N.A., Md Sabri, B.A.: The relationship of educational environment and preparedness to practice-perceptions of Malaysian dental graduates. Eur. J. Dent. Educ. **27**(3), 449–456 (2022)
11. Puth, M.T., Neuhäuser, M., Ruxton, G.D.: Effective use of spearman's and kendall's correlation coefficients for association between two measured traits. Anim. Behav. **102**, 77–84 (2015)
12. Shongwe, M., Ocholla, D.N.: A tracer study of LIS graduates at the university of zululand, 2000–2009. Mousaion **29**(2), 227–245 (2011)

Hardware Design and Implementation of a Low-Cost IoT-Based Fire Detection System Prototype Using Fuzzy Application Methods

Emmanuel Lule[1,2(✉)], Chomora Mikeka[3], Alexander Ngenzi[1], Didacienne Mukanyiligira[4], and Parworth Musdalifah[2]

[1] African Center of Excellence in IoT, University of Rwanda, Nyarugenge 3900, Rwanda
[2] Department of Computer Science, Makerere University, Kampala 7092, Uganda
emmanuel.lule@mak.ac.ug
[3] Directorate of Science, Technology and Innovations, Min. of Education, Lilongwe 328, Malawi
[4] National Council for Science and Technology (NCST), Kigali 2285, Rwanda

Abstract. The lack of reliable, low-cost fire detection systems within the central local market communities of East Africa (EA) has resulted in serious catastrophic fire accidents that have resulted in loss of life and property. Proposed satellite systems are expensive to acquire and maintain. Also, unit sensor smoke detectors are highly susceptible to false alarm rates. The proposed low-cost fire detection system prototype was implemented using the Arduino UNO, Arduino IDE platforms integrated with the fuzzy logic technique to determine an informed fire status decision. Hence, the proposed solution ensures appropriate public fire safety protection for the nearby market community by providing early warning alarm notification in case of a fire outbreak. Obtained results show that the proposed fire detection system prototype using fuzzy logic exhibited an accuracy rate of 91% using a confusion matrix model (CMM) evaluation method.

Keywords: Internet of Things (IoT) · Fuzzy Application Methods · Embedded Sensors and Confusion Matrix Model (CMM)

1 Introduction

1.1 Background Study

Local urban markets located within metropolitan areas of East Africa (EA), are faced with repeated fire accidents. For example; Gisozi, Rwanda; St. Balikudembe (Owino), Uganda; and Gikomba, Kenya. From the investigations conducted, little or no effort has been made to ensure sufficient control and suppression of fires for the purposes of safeguarding and protecting the nearby communities. In Uganda, the major causes of fires are attributed to electrical short circuits, suspected arson, neglected charcoal stoves, and negligence, among others [1–4]. Because of the rampant, uncontrolled fires, this has resulted in the destruction of properties and human life.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 61–76, 2024.
https://doi.org/10.1007/978-3-031-53552-9_6

Figure 1 shows the number of victims affected (injured or fatal) by fire accidents in Uganda spanning a period of 2012–2020. Guohua *et al.* [5, 6] propose a unit smoke detector for forest fire detection using deep learning. However, such systems are prone to false alarms due to their high sensitivity, rendering them inaccurate for fire detection. Mazzeo *et al.* [7] propose a satellite based system for effective fire detection, however, they are expensive to acquire and maintain for developing countries. Also, satellites incur unnecessary delays due to long scanning cycles [8]. Enormous research in fire detection using fuzzy logic has been done. Surya *et al.* [9], developed a smart home fire detection system using fuzzy logic. Raspberry PI, DHT11, and MQ2 sensors were used to detect fires. Results show a notification through WhatsApp, and monitoring is achieved via a web application interface.

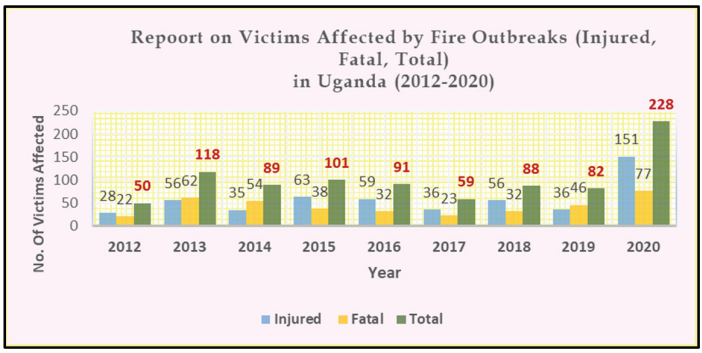


Fig. 1. Injured, Fatal and Total No. of Victims Affected by Fire Accidents, in Uganda Between the Period of (2012–2020) [3, 10]

2 Related Works

Numerous studies have been conducted in the fire detection area. For instance, Shereiqi *at el.* [11] proposes an IoT alarm system for smart buildings with smoke and temperature sensors. Notification messages are sent to the security team when fire is detected, along with the time and location. The paper proposes a low-cost, low power multi-sensory fire detection system to compensate for the unit-smoke detectors and reduce false alarms. Ivan *et al.* [12] from Serbia, use the index and fuzzy AHP methods together with the TOPOSIS method for forest fire susceptibility zonation integrated with IoT. Results indicate very high and high forest susceptibility indexes of 26.85% and 25.75% for the fuzzy Index zone, respectively. The method is used in forest fire risk assessment to improve monitoring and early warning mechanisms.

Stavros *et al.* [13] implore spatiotemporal analysis to determine the severity of fire hazards. Two modeling techniques, the Analytical Hierarchy Process (AHP), and Fuzzy AHP, are used to estimate fire hazards in 20 years. Geostatistical analysis revealed a significant clustering process of high-risk values in the southwest and northern parts of the study area, and clusters of low-risk values are clustered in the northern territory.

The degree of spatial autocorrection tends to be greater for 1996 than 2016, with higher fire hazard transmission risk in most regions in the past. Rehman *et al.* [14], proposed a system of sensors for temperature, humidity, heat, and smoke to detect fire. Using an AI-based fuzzy algorithm with specific rule sets, an informed decision about fire status is determined. The system provides alerts and hardware control mechanisms, i.e., an open ventilation system in case of suffocation, or initiating a water sprinkler. Results show 15 performance tests between different fire intensities using MATLAB.

Vasanthkurmar *et al.*[15], propose a GSM-IoT based firefighting robot. It uses an IR sensor for fire detection using the nozzle head of firefighting robots and sends information to the MCU programmed with a fuzzy algorithm. Nebot [16] proposes a hybrid Inductive Fuzzy Reasoning (FIR) and Neural Fuzzy Inference System (ANFIS) to model burned areas in Portugal's forests. Results were accurate compared to other AI techniques. Bhuvanesh *et al.* [17], propose a network of WSNs placed over the forest to detect fire. Multiple sensors for temperature, humidity, light, and gas are used. Integrated with the data fusion techniques of type 2 fuzzy systems. Results enhance the consistency and exactness of true fire incidents.

Ren *et al.* [18] suggest an intelligent detection technology for fires using multi-information fusion and fault detection methods for green buildings. Using fuzzy reasoning, the extracted information fusion is used to detect arc faults that cause electrical fires in the low distribution systems of green buildings. The results were satisfactory. Hence, this paper, presents a hardware design and implementation prototype of a low-cost, low-power IoT-based fire detection system using fuzzy methods for ensuring early fire detection in local markets by making an appropriate and informed decision. The developed fire detection system is also easily deployable and reproducible for supporting early fire detection in markets with the aim of ensuring public fire safety and protection within the vendor communities by promoting safe and safe evacuation of persons.

3 Methodology

The paper considers an experimental hardware implementation prototype of a fire detection system using fuzzy application methods. The hardware prototype uses an embedded systems Arduino UNO board integrated with the Arduino integrated development environment (IDE) software (Ver. 1.8.15). Various sensors are interfaced with the Arduino UNO through the breadboard. Sensor modules, i.e., DHT11: for temperature and humidity, MQ2; for smoke detection, MQ135; for detecting CO2 dissipated in the atmosphere due to a fire outbreak. The flame sensor detects IR flames with configurations of KY-026, or LM 393. If flame = 0, "fire is detected" or flame = 1, "no fire". The Expt. is interfaced with the actuator sensors, i.e., a buzzer for alarm signaling when fire is detected and an LED for light signaling when fire is detected. The "Red-LED" represents fire detected and the "Green-LED" for smoke detected in the proposed Expt. (cf. Figs. 4 and 6). Then, the ESP8266 module provides internet communication to send all the collected data to the Thing Speak cloud API for insightful analysis purposes.

4 Fuzzy Application Methods

Fuzzy Logic is based on "degree of truth" rather than Boolean Logic for "True or False" (1 or 0) in the concept of modern computing [19]. This paper considers temperature, humidity, smoke, and CO2 gas as the crisp inputs for fire detection (cf. Fig. 2). We utilize the "***fuzzification***" method to convert crisp inputs into fuzzy values with the defined knowledge base (cf. Fig. 2) [1, 20]. Thus, the defined fuzzy membership values corresponding to Temperature and CO2 input parameters are $\Delta T = \{\text{Low, Moderate, High}\}$, and $CO2 = \{\text{Very Low, Low, Moderate, High, Very High}\}$. While, "***defuzzification***" is the process where output fuzzy sets are reconverted back to a single crisp number logic. The output Fire Index (FI) defines fire intensity output per unit time for a given area, converted into a fuzzy range of [0–1]. Hence, $FI = \{\text{VL, L, M, H, and VH}\}$. We proposed five (5) of Mamdani's sparse knowledge inference rules to define the most significant operations of the fire detection system prototype. Mamdani's inference system are widely acceptable for the development of fuzzy expert system applications with rules created from human intelligence knowledge [21]. Secondly, to minimize the system complexity of the fire detection prototype, we utilized few significant inference rules. The sparse fuzzy inference system (FIS) rules were defined as follows:

1. IF temp is **Low** AND CO2 is **Very Low** THEN FI is **VL**
2. IF temp is **Low** AND CO2 is **Low** THEN FI is **L**
3. IF temp is **Moderate** AND CO2 is **Moderate** THEN FI is **M**
4. IF temp is **High** AND CO2 is **High** THEN FI is **H**
5. IF temp is **High** AND CO2 is **Very High** THEN FI is **VH**

The fuzzy inference system (FIS) design, forms a corresponding crisp output design through defuzzification. The output is obtained using the common defuzzification method "***Centroid***". This method determines the center of the area of fuzzy sets and then returns the corresponding crisp value [22–24].

4.1 Derived Mathematical Fuzzy-Based Membership Functions (MF) of the Proposed Fire Detection System Prototype

The membership function (MF) fuzzy set information for Temperature, Carbon dioxide (CO2), and the output Fire Index or Intensity (FI) is represented graphically as follows:

The derived mathematical fuzzy membership functions (MF) of element x in set X, the degree of membership $\mu X[x]$ to the unit interval [0–1] in Fig. 2a, for input temperature using the triangular MF can be explicitly defined using the fuzzy set equations from Eqs. 1, 2 and 3 below:

$$\mu\text{Low}[x] = \begin{cases} 1, x \leq 0 \\ \frac{x-0}{20-0}, 0 \leq x \leq 20 \\ \frac{40-x}{40-20}, 20 \leq x \leq 40 \\ 0, 40 \leq x \end{cases} \quad (1)$$

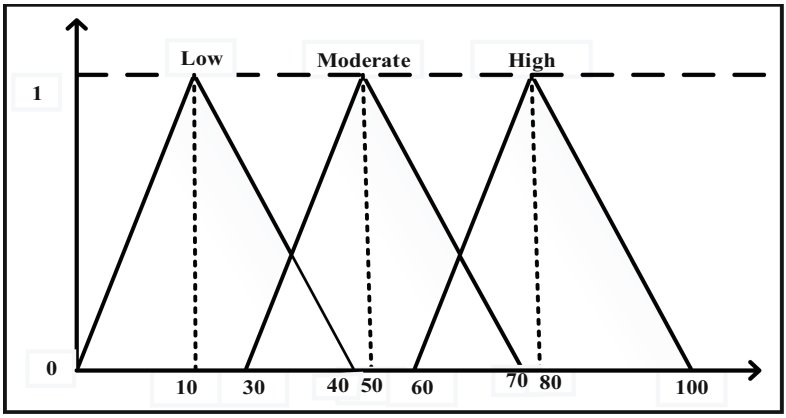


Fig. 2a. Parameter Input Temperature Membership Function (MF) with Fuzzy Set Variations.

$$\mu\text{Moderate}[x] = \begin{cases} 1, x \leq 30 \\ \frac{x-30}{50-30}, 30 \leq x \leq 50 \\ \frac{70-x}{70-50}, 50 \leq x \leq 70 \\ 0, 40 \leq x \end{cases} \tag{2}$$

$$\mu\text{High}[x] = \begin{cases} 1, \leq 60 \\ \frac{x-60}{80-60}, 0 \leq x \leq 20 \\ \frac{100-x}{100-80}, 20 \leq x \leq 40 \\ 0, 100 \leq x \end{cases} \tag{3}$$

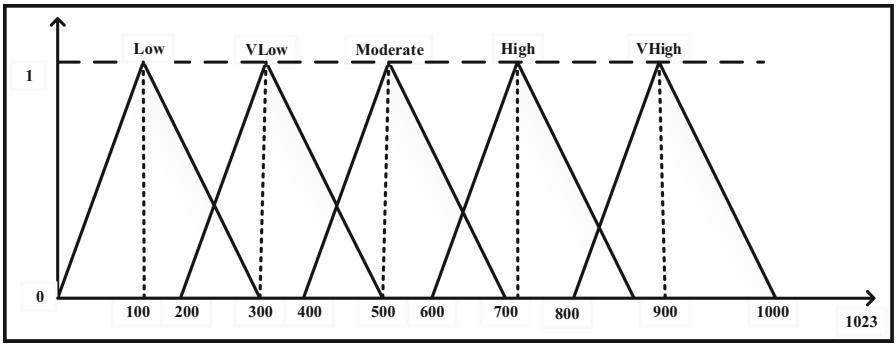


Fig. 2b. Carbon dioxide (as Input) MF Fuzzy Set Variations.

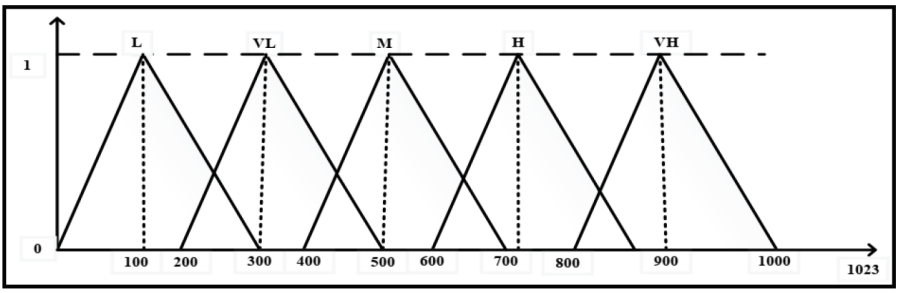


Fig. 2c. Fire Index (as Output) MF Fuzzy Set Variations.

Likewise, the mathematical fuzzy MF for input Carbon dioxide (cf. Fig. 2d) and output Fire Intensity (FI) (cf. Fig. 2c), can be implicitly defined as follows: (cf. Eqs. 4, 5, 6, 7 and 8).

$$\mu\text{VLow}[x] \text{ or } \mu\text{VL}[x] = \begin{cases} 1, x \leq 0 \\ \frac{x-100}{100-0}, 0 \leq x \leq 100 \\ \frac{300-x}{300-100}, 100 \leq x \leq 300 \\ 0, 300 \leq x \end{cases} \tag{4}$$

$$\mu\text{VLow}[x] \text{or } \mu\text{VL}[x] = \begin{cases} 1, x \leq 200 \\ \frac{x-200}{300-200}, 200 \leq x \leq 300 \\ \frac{400-x}{400-300}, 300 \leq x \leq 400 \\ 0, 400 \leq x \end{cases} \tag{5}$$

$$\mu\text{Moderate}[x] \text{or } \mu\text{M}[x] = \begin{cases} 1, x \leq 300 \\ \frac{x-300}{500-300}, 300 \leq x \leq 500 \\ \frac{500-x}{700-500}, 500 \leq x \leq 700 \\ 0, 700 \leq x \end{cases} \tag{6}$$

$$\mu\text{High}[x] \text{or } \mu\text{H}[x] = \begin{cases} 1, x \leq 600 \\ \frac{x-600}{700-300}, 600 \leq x \leq 700 \\ \frac{800-x}{800-700}, 700 \leq x \leq 800 \\ 0, 800 \leq x \end{cases} \tag{7}$$

$$\mu\text{High}[x] \text{or } \mu\text{H}[x] = \begin{cases} 1, x \leq 750 \\ \frac{x-750}{900-750}, 750 \leq x \leq 900 \\ \frac{1023-x}{1023-900}, 900 \leq x \leq 1023 \\ 0, 1023 \leq x \end{cases} \tag{8}$$

4.2 Proposed Fuzzy Logic Controller for the Low-Cost Fire Detection System

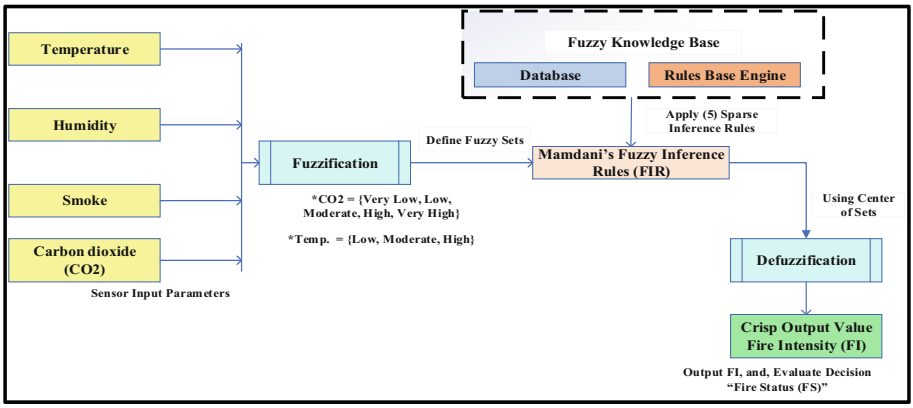


Fig. 2d. Proposed Fuzzy Based Controller for the Low-Cost Hardware-Based Fire Detection System.

4.3 Architectural Overview of the Proposed Fuzzy-Based Fire Detection System Prototype

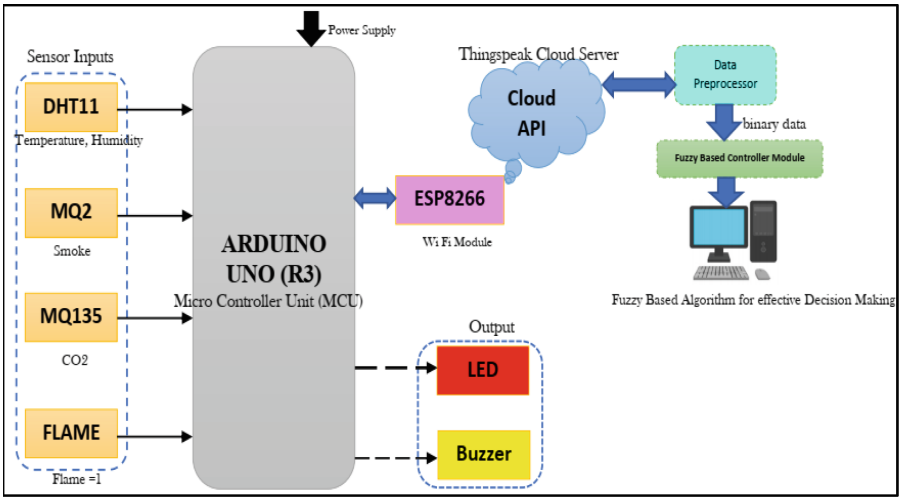


Fig. 3. Architectural Overview of the Proposed Fuzzy-Based Fire Detection System Prototype.

Figure 3 shows the architectural design of the hardware fire detection prototype. The different components interface with the Arduino UNO microcontroller unit (MCU). Four sensor modules are required to collect data about the fire outbreak. E.g., DHT11, MQ2,

MQ135 and Flame (LM-393) or (KY-026). Sensors collect data regarding; temperature, humidity, CO_2, smoke, and flame. Data is sent via the ESP8266 Wi-Fi module to the open-source Thing Speak cloud API platform for insightful analysis, which is stored as a "*.csv" file, an equivalent of a data pre-processor. Two actuator sensor modules, the Buzzer and LED, are interfaced for alerting the authorities through a signal notification in case of a fire outbreak through a sound alarm (Buzzer) or a light warning signal (LED). Using the fuzzy method, ΔT, and CO_2 data are analyzed and evaluated to make an informed decision about the "*Fire Status*". Figure 2d shows the fuzzy controller for the proposed fire detection prototype. Having Temperature (ΔT), and gas (CO_2) as the primary parameters to be fuzzified, i.e., CO_2 = {VLow, Low, Moderate, High, VHigh}; Temperature = {Low, Moderate, High}, (cf. Fig. 2a-b). The knowledge base containing the determined sparse fuzzy inference rules (FIR), is then applied to the proposed fire detection prototype. Likewise, a defuzzification process is successfully used to determine the crisp output value called "Fire Intensity" or "Fire Index (FI)", of the system (cf. Fig. 2c). Note that, the fuzzy membership set is defined as: FI = {VL, L, M, H, and VH}. Hence, fire intensity is the rate of heat transfer per unit time.

5 Schematic Circuit Design of the Proposed Fire Detection System Prototype

Figure 4 shows the schematic circuit connection diagram of the proposed fire detection system prototype using Fritzing (ver. 0.9.3b) software. An open-source software that supports the design of the hardware-based electronic printed circuit boards (PCB) to be

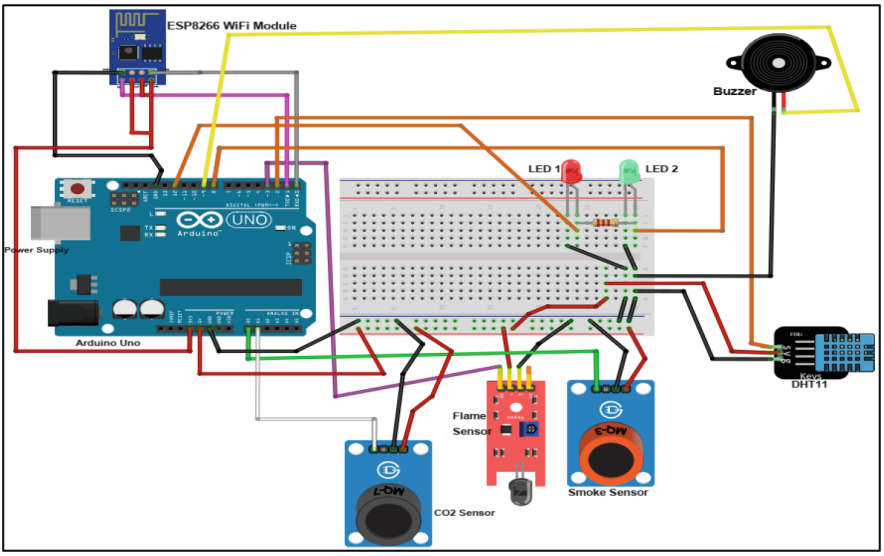


Fig. 4. Schematic Diagram of the Proposed Fire Detection System Prototype Using Fritzing Software.

used by scientists, and engineers before building and implementing IoT-based hardware prototypes (Tables 1 and 2).

5.1 Hardware Requirement

Table 1. Hardware Requirements Considered

Hardware	Parameter	Purpose/Description
DHT 11	Temperature, Humidity	Detect Temp, Humidity
Flame Sensor (LM-393)	Flame	Detects presence of flame
MQ2	Smoke	Detects Presence of Smoke
MQ135	CO_2	Detects Presence of CO_2
ESP8266 Wi-Fi Module	Internet Connection through MQTT protocol	Send Data to the Cloud API via Internet
LED	Light Signaling	Provide Light Warning
Piezo Buzzer	Sound alert notification	Provide Sound Alert
Resistor	220Ω Resistor	To drop the Vcc current that is not required to operator the LED

5.2 Software Requirement

Table 2. Software Requirements Considered

Software	Purpose
Fritzing Software (Ver. 0.9.3b)	Designing of electronic circuits before production
Arduino IDE (Ver. 1.8.15)	Provide a programming environment for the development of embedded system hardware devices
Thing Speak Cloud API	IoT Cloud Server Platform for data storage and analysis

6 Framework of the Proposed Fire Detection System Prototype Using Fuzzy Methods

Figure 5 represents the framework of the fire detection prototype using a flowchart. The solution accepts temperature, humidity, smoke, CO2 and flame as input parameters to the system. If an infra-red (IR) flame is detected, then “fire is detected”, otherwise “no fire

detected". A signal notification is then sent to authorities for possible warning, once fire is detected. Also, when "Smoke > 230" ppm, is detected, authorities are notified through an alarm notification for immediate action. Collected data is sent to the Thing Speak cloud API for storage and analysis to draw useful insights. Mamdani's fuzzy method is then applied to two inputs, i.e., CO2, and Temperature (ΔT) as primary factors contributing to a fire outbreak noticed by a certain change from normal conditions. Through the "***fuzzification***" and "***defuzzification***" processes, we obtain the "*Fire Intensity*" or "***Fire Index** (FI)*" as the output in the range of [0–1] corresponding to the fuzzy decision "***Fire Status** (FS)*" determined. The process is continued until the value of CO2 is below the threshold value of 1000 ppm, otherwise, it is terminated.

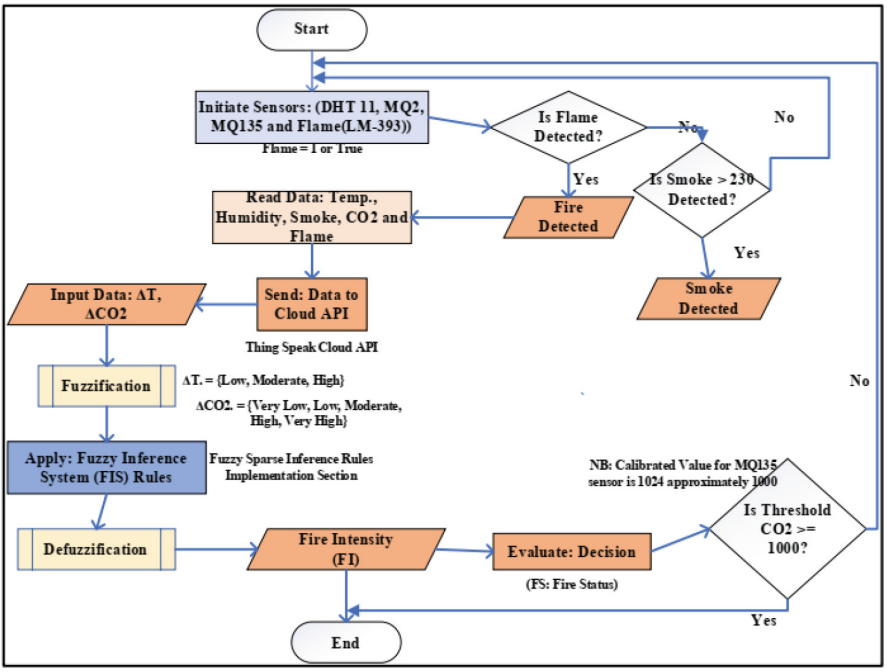


Fig. 5. The Proposed Framework Workflow of the Fire Detection System Prototype Using Fuzzy Application Methods.

6.1 Lab. Experimental Setup

In Fig. 6, we show a laboratory experimental setup of the proposed IoT-enabled low-cost fire detection prototype using fuzzy application methods. The experiment consists of several embedded system sensor modules, namely; DHT11 for measuring temperature and humidity; MQ2 for smoke detection; MQ135 for detection of atmospheric carbon dioxide in the surrounding area; a buzzer for alarm notification; and an LED for light signaling. The proposed components have been connected to the Arduino UNO board by

using jumper wires. A 220Ω resistor is connected to regulate the flow of current through the circuit. Datasets of humidity, temperature, smoke, and CO_2 are collected through the Thing Speak cloud API. The ESP8266 Module provides internet communication. Using a burning candle, we carefully simulated a fire outbreak using a candle flame specimen in a lab. Setting to demonstrate a fire event situation. Following the simulated fire event, results were obtained, which can be viewed on the output serial monitor screen (cf. Fig. 7). A further summary of the results from the serial monitor output is tabulated in (cf. Table 3).

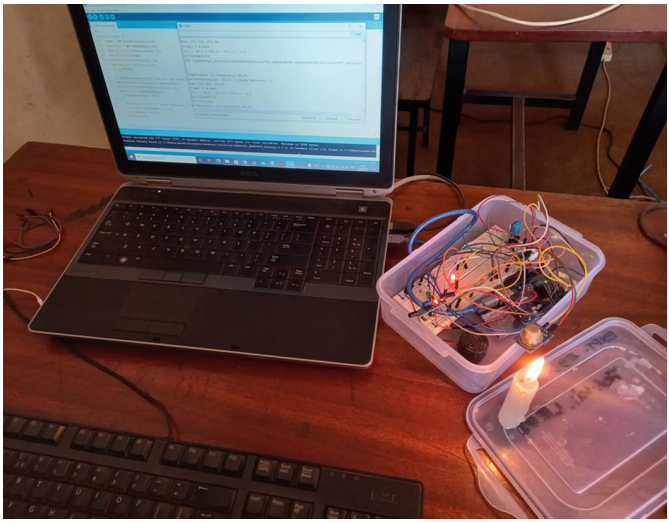


Fig. 6. Lab. Experimental Setup for the Hardware Based Low-Cost Fire Detection System Prototype.

6.2 Serial Monitor Output

Figure 7 shows a detailed snapshot of the captured output of the serial monitor results from the Lab. Experiment. The left-hand side of Fig. 7 shows the fire detection code developed using the Arduino IDE platform. The primary source of input raw data collected from the sensor readings includes: temperature (^{0}C), humidity (%), smoke level measured in parts per million (ppm), flame sensor (1 = True, 0 = False); smoke Condition detected when (Smoke > 230); Fire Condition, Fuzzy input for the temperature sensor [0–100]; and Fuzzy input for CO_2 [1–1023] for an MQ135 calibrated sensor; the associated "Fire index (FI)"; and prevailing "Fire Status (FS)" are fully represented below (cf. Table 3).

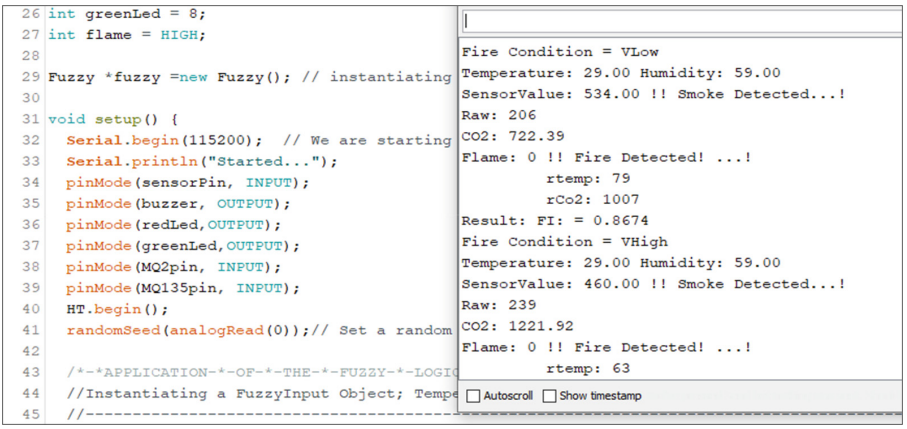


Fig. 7. Sampled Snapshot of the Serial Monitor Output for the Proposed Fire Detection System Source Code.

Table 3. Summary of the Extracted Results Output of the Proposed Low-Cost Fire Detection Systems Lab. Expt.

Expt. No	Temp. (^{0}C)	Hum. (%)	Smoke conc	Smoke cond	CO_2 conc	CO_2 (ppm)	Flame sensor	Fire cond	Fuzzy temp	Fuzzy CO_2	FI	FS
1	29	59	225	No Smoke	138	195.38	1	No Fire	72	59	0.0000	VLow
2	29	59	227	No Smoke	141	204.55	1	No Fire	58	49	0.0000	VLow
3	29	59	227	No Smoke	141	199.93	1	No Fire	66	545	0.4888	Medium
4	29	59	228	No Smoke	142	209.25	1	No Fire	21	45	0.1367	VLow
5	28	58	221	No Smoke	131	161.72	1	No Fire	59	277	0.0000	VLow
6	29	59	220	No Smoke	132	165.68	1	No Fire	41	758	0.0000	VLow
7	28	58	220	No Smoke	132	165.68	1	No Fire	76	460	0.0000	VLow
8	29	59	227	No Smoke	143	214.02	1	No Fire	31	500	0.4888	Medium
9	28	58	243	No Smoke	160	308.32	1	No Fire	59	565	0.4888	Medium
10	28	58	228	No Smoke	140	199.93	1	No Fire	0	657	0.0000	VLow
11	28	58	230	No Smoke	141	204.55	0	!!Fire Detected!!	15	1006	0.0000	VLow
12	28	58	232	No Smoke	144	218.88	0	!!Fire Detected!!	65	701	0.6843	High
13	28	58	232	No Smoke	143	214.02	0	!!Fire Detected!!	61	589	0.4888	Medium
14	28	58	232	No Smoke	143	214.02	0	!!Fire Detected!!	67	936	0.8687	VHigh
15	28	58	232	No Smoke	143	214.02	1	No Fire	10	203	0.1442	VLow
16	28	58	257	No Smoke	140	199.93	1	No Fire	49	881	0.0000	VLow

(*continued*)

Table 3. (*continued*)

Expt. No	Temp. (^{0}C)	Hum. (%)	Smoke conc	Smoke cond	CO_2 conc	CO_2 (ppm)	Flame sensor	Fire cond	Fuzzy temp	Fuzzy CO_2	FI	FS
17	29	59	517	!!Smoke Detected!!	150	255.22	1	No Fire	23	745	0.0000	VLow
18	28	58	420	!!Smoke Detected!!	160	321.17	1	No Fire	75	696	0.6843	High
19	29	59	374	!!Smoke Detected!!	139	199.93	1	No Fire	42	358	0.4888	Medium
20	29	59	358	!!Smoke Detected!!	149	239.14	0	!!Fire Detected!!	75	181	0.0000	VLow
21	29	59	346	!!Smoke Detected!!	154	277.89	0	!!Fire Detected!!	41	550	0.4888	Medium
22	29	59	349	!!Smoke Detected!!	155	283.79	0	No Fire	61	858	0.8669	VHigh

KEY: VLow = Very Low; **VHigh** = Very High; **FI** = Fire Index; **FS** = Fire Status; **ppm** = parts per million; **Smoke Cond.** = Smoke Condition; **CO_2 Conc.** = CO_2 Concentration; **Fuzzy Temp.** = Fuzzy Input for Temperature in range of [1–100]; and **Fuzzy CO_2** = Fuzzy Input for CO_2 in range of [1–1023].

6.3 Extracted Summary of Obtained Result Output from Lab. Experiment

7 Results and Discussion

In Figs. 8, 9, 10, 11, 12 and 13 we show several insights derived from the Lab. Experiment using Thing Speak and MATLAB visualization tools. A correlation between temperature, humidity, smoke, and CO_2 is presented and discussed below.

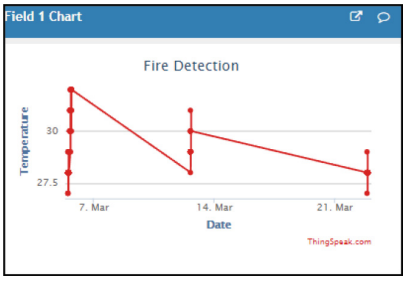


Fig. 8. Temp. Variation with Day/Time

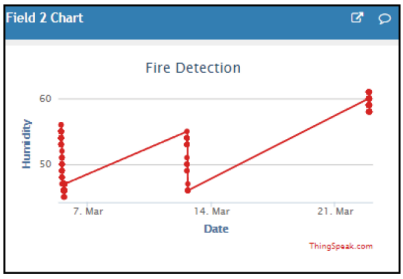


Fig. 9. Humidity Variation with Day/Time

Figure 8 shows the temperature variation with day and time, which gradually increases once a "***fire is detected***" above 300 °C. This is then followed by a subsequent decline in temperature values back to normal for "***no fire detected***." Likewise, increased temperatures significantly decrease the humidity levels of the atmospheric surroundings. In Figs. 10 and 11, the variation of smoke vs. CO_2 is represented. The results concluded that the concentration of smoke in the atmosphere increases once the smoke particles are detected. Hence, fire detection significantly increases the level of carbon dioxide (CO_2) in the atmosphere above the threshold value of 1000 ppm.

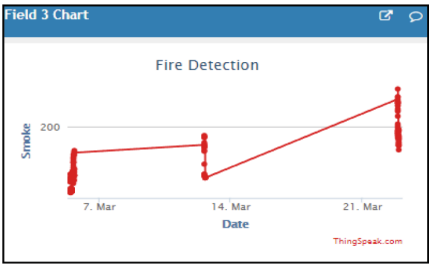


Fig. 10. Smoke Variation with Day/Time

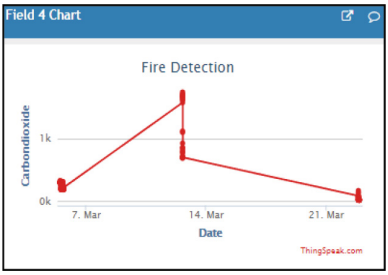


Fig. 11. CO_2 Variation with Day/Time

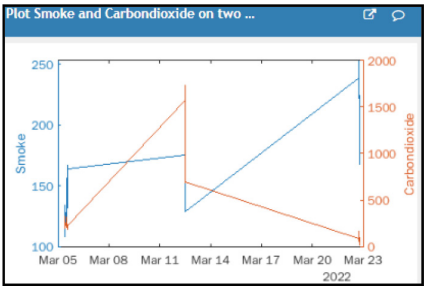


Fig. 12. Smoke Vs CO_2

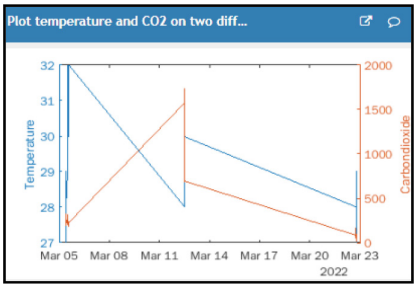


Fig. 13. Temp. Vs CO_2 Using MATLAB.

Figures 12 and 13 show a comparison analysis of temperature, smoke, and CO_2 as input parameters by using the Thing Speak cloud API integrated with the MATLAB Visualization Tool. Figure 12 shows a gradual increase in smoke concentration, which increases the level of CO_2, followed by subsequent drops. Figure 13 demonstrated that increased temperatures necessitate a gradual increase in the level of carbon dioxide dissipated above the threshold value, followed by sudden drops at reduced temperatures.

8 Performance Evaluation of the Proposed Fire Detection System Prototype

In Table 4, we compute the operating accuracy rate of the proposed low-cost fire detection prototype by classifying the obtained values of the output "**Fire Status (FS)**" from the obtained "**Fire Index** (FI)", in the range of [0–1] (cf. Table 3) to successfully classify the **True Positive (TP), False Positive** (FP), and **False Negatives** (FN) and **True Negative** (TN) using the Confusion Matrix Model (CMM).

The selected sample contained 22 datasets (cf. Table 3) to predict the "true positives" or "true negatives" positives" of the rule outcomes. The confusion matrix defines a summary of the prediction outcome or results of a typical fuzzy based classification

problem. Thus, the accuracy rate of the Lab. Experiment is equivalent to 91% (cf. Eq. 9).

$$\frac{TP+TN}{TP+TN+FP+FN} \text{ x } 100\% \tag{9}$$

Table 4. The Confusion Matrix Model (CMM) for the Proposed Low-Cost Fire Detection System Using Fuzzy Application Methods

Sampled dataset size (N = 22)	Observed positive values	Observed negative values	Total
Predicted Positive Values	TP = 19	FP = 2	21
Predicted Negative Values	FN = 0	TN = 1	01

9 Conclusion and Future Works

The study presents a novel idea of using Mamdani's' sparse-based fuzzy inference application methods in the design and implementation of a hardware prototype for a low-cost fire detection system for local urban markets in the East Africa (EA), with a major purpose of ensuring early fire protection and safety within the community. The experimental results obtained (sample size N = 22 rule outcomes) in Table 3 were evaluated using the confusion matrix model (CMM), which achieved an operating accuracy rate of 91%. Hence, the proposed solution shall assist the fire and rescue department, the local vendor community as a foundation in providing reliable early warning notifications by promoting appropriate public fire safety and protection measures through ensuring quick evacuation of affected persons. Future works intend to utilize machine learning (ML) or convolutional neural networks (CNN) approaches to design and implement effective low-cost fire detection systems or devices to significantly increase the rate of accuracy of the proposed solution prototype, thus minimizing the rate of false alerts.

Acknowledgement. Thanks to Prof. Chomora Mikeka for Providing all the technical information leading to the article. Thanks to Dr. Alexander Ngenzi and Dr. Didacienne Mukanyiligira for contributing to the knowledge in the drafting and conceptualization of the conference paper article.

References

1. Lule, E., Mikeka, C., Ngenzi, A., Mukanyiligira, D.: Design of an IoT-based fuzzy approximation prediction model for early fire detection to aid public safety and control in the local urban markets. Symmetry (Basel). **12**(9), 1391 (2020)
2. Lule, E., Eddie Bulega, T.: A scalable wireless sensor network (WSN) based architecture for fire disaster monitoring in the developing world. Int. J. Comput. Netw. Inf. Secur. **7**(2), 40–49 (2015)

3. UPF: Uganda Police Annual Crime Report., Kampala (2021)
4. Uganda Police: Annual Crime Report. 184 (2020)
5. Wang, G., Li, J., Zheng, Y., Long, Q., Gu, W.: Forest smoke detection based on deep learning and background modeling. In: Proceeding of 2020 IEEE International. Conference in Power, Intelligent Computing and Systems. (ICPICS), Syenyang, China, 28–30th July 2020
6. Nguyen, H.: A fuzzy-based smoke detection on embedded system. J. Theor. Appl. Inf. Technol. **97**(12), 3415–3424 (2019)
7. Mazzeo, G., et al.: Integrated satellite system for fire detection and prioritization. Remote Sens. **14**(2), 1–25 (2022)
8. Bhattacharya, S., et al.: Experimental analysis of WSN based solution for early forest fire detection. In: Proceeding of 2021 IEEE International Conference in Internet Things and Intelligent Systems (IoTaIS), Bandung, Indonesia, 23–24th November 2021
9. Surya Devi, A.A.P.B., Istikmal, K.N.: Design and implementation of fire detection system using fuzzy logic algorithm. In: Proceeding of 2019 IEEE Asia Pacific Conference in Wireless and Mobile (APWiMob), Bali, Indonesia, 5–7th November 2019
10. Uganda Police: Uganda Police Annual Crime Report (2022)
11. Al Shereiqi, I.M., Sohail, M. mad: smart fire alarm system using IOT. J. Student Res. 1–9 (2020)
12. Novkovic, I., et al.: Gis-based forest fire susceptibility zonation with IoT sensor network support, case study—nature park Golija. Serbia. Sens. **21**(19), 1–29 (2021)
13. Sakellariou, S., et al.: Remotely sensed data fusion for spatiotemporal geostatistical analysis of forest fire hazard. Sensors (Switzerland). **20**(17), 1–20 (2020)
14. Rehman, A., et al.: Smart fire detection and deterrent system for human savior by using internet of things (IoT). Energies. **14**(17), 5500 (2021)
15. Vasanthkumar, P., Arunraj, P. V., Khan, N.M.B., Akash, A.V., Mukunthan, R., Babu, R.H.: Fuzzy logic algorithm and GSM IoT based fire fighting robot. J. Phys. Conf. Ser. **2040**(1), 012045 (2021)
16. Nebot, À., Mugica, F.: Forest fire forecasting using fuzzy logic models. Forests. **12**(8), 1005 (2021)
17. Bhuvanesh, A., Kannan, S., Babu, M.A., Rose, J.L., Dhanalakshmi, S.: Application of interval type-2 fuzzy logic data fusion using multiple sensors to detect wildfire. Tierärztliche Prax. **41**, 530–539 (2021)
18. Ren, X., et al.: Design of multi-information fusion based intelligent electrical fire detection system for green buildings. Sustain. **13**(6), 3405 (2021)
19. Syafitri, N., et al.: Early detection of fire hazard using fuzzy logic approach. Int. J. Adv. Comput. Res. **9**(43), 252–259 (2019)
20. Rachman, F.Z., et al.: Design of the early fire detection based fuzzy logic using multisensor. IOP Conf. Ser. Mater. Sci. Eng. **732**(1), 012039 (2020)
21. Kushnir, A., Kopchak, B., Oksentyuk, V.: Development of heat detector based on fuzzy logic using arduino board microcontroller. In: Proceeding OG 2023 IEEE 17th International Conference on the Experience of Designing and Application of CAD Systems, Jaroslaw, Porland, 22–25th Febuary 2023
22. Labellapansa, A., et al.: Prototype for early detection of fire hazards using fuzzy logic approach and Arduino microcontroller. Int. J. Adv. Comput. Res. **9**(44), 276–282 (2019)
23. Saponara, S., Elhanashi, A., Gagliardi, A.: Real-time video fire/smoke detection based on CNN in antifire surveillance systems. J. Real-Time Image Process. **18**(3), 889–900 (2021)
24. Gaur, A., et al.: Fire sensing technologies: a review. IEEE Sens. J. **19**(9), 3191–3202 (2019)

Closing the Data Gap: A Comparative Study of Missing Value Imputation Algorithms in Time Series Datasets

Sepideh Hassankhani Dolatabadi[1](✉), Ivana Budinská[1], Rafe Behmaneshpour[2], and Emil Gatial[1]

[1] Institute of Informatics, Slovak Academy of Sciences, Bratislava, Slovakia
{sepideh.ui,Ivana.Budinska,emil.gatial}@savba.sk
[2] Rahbar Farayand Arya Company, Department of Maintenance, Bratislava, Slovakia

Abstract. The presence of missing values in time series datasets poses significant challenges for accurate data analysis and modeling. In this paper, we present a comparative study of missing value imputation algorithms applied to time series datasets collected from various sensors over a period of six months. The goal of this study is to bridge the data gap by effectively replacing missing values and assessing the performance of three common imputation algorithms for time series: K-Nearest Neighbors (KNN) imputer, Expectation-Maximization (EM), and Multiple Imputation by Chained Equations (MICE). To evaluate the performance of the imputation techniques, we employed Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) as metrics. Through rigorous experimentation and analysis, we found that each algorithm exhibited varying degrees of effectiveness in handling missing values within the time series datasets. Our findings highlight the importance of choosing an appropriate imputation algorithm based on the characteristics of the dataset and the specific requirements of the analysis. The results also demonstrate the potential of the MICE imputer in closing the data gap and improving the accuracy of subsequent analyses on time series sensor data. Overall, this study provides valuable insights into the performance and suitability of different missing value imputation algorithms for time series datasets, facilitating better decision-making and enhancing the reliability of data-driven applications in various domains.

Keywords: Missing value imputation · Data preprocessing · Sensor data analysis

Abbreviations

The following abbreviations are used in this manuscript:

KNN	K Neareast-Neighbour
EM	Expectation Maximization
MICE	Multiple Imputation by Chained Equations
RMSE	Root Mean Squared Error
MAE	Mean Absolute Error

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 77–90, 2024.
https://doi.org/10.1007/978-3-031-53552-9_7

MVI	Missing Value Imputation
EMMVI	Expectation-Maximization Missing Value Imputation
MMVI	Multiple Imputation by Chained Equations (MICE)
LLSMVI	Locally Linear Stochastic Missing Value Imputation
BPCAMVI	Bayesian Principal Component Analysis Missing Value Imputation
WSNs	Wireless sensor networks
LRMVI	Latent Regression Missing Value Imputation
NRMSE	Normalized Root Mean Squared Error
MSE	Mean Squared Error
RF	Random Forest
SVM	Support Vector Machines
BPCA	Bayesian Principal Component Analysis
DT	Decision Tree
ML	Machine Learning
CVBKNNI	Cross-Validation Based k-Nearest Neighbor Imputation
RNNs	Recurrent Neural Networks
MuSDRI	Multi-Seasonal Decomposition based Recurrent Imputation
HPGR	High-Pressure Grinding Rolls
IQR	Interquartile Range

1 Introduction

Missing values pose a significant challenge in data analysis and can have a substantial impact on the reliability and validity of the findings. In time series datasets, where observations are collected over successive time points, missing values are particularly prevalent and can disrupt the temporal patterns and dependencies inherent in the data. Dealing with missing values is crucial for accurate analysis and meaningful interpretation of time series data. Missing values can occur in time series datasets due to various reasons, including equipment malfunctions, data transmission errors, and incomplete data collection processes. The presence of missing values introduces gaps in the time series, which can disrupt the continuity of the data and hinder the extraction of valuable insights. Therefore, it is essential to address missing values appropriately to avoid biased or erroneous conclusions.

In recent years, researchers have devoted significant attention to developing effective techniques for handling missing values in time series datasets. These techniques aim to estimate or impute the missing values based on the available data, thereby preserving the temporal structure and ensuring the completeness of the time series [1]. Several strategies have been proposed and employed in practice to address the issue of missing values in time series analysis. One common approach for handling missing values in time series datasets is imputation. Imputation involves filling in the missing values with estimated values based on the observed data. Numerous imputation methods have been proposed, each with its own underlying assumptions and algorithms. Some commonly used imputation techniques include mean imputation, forward/backward filling, linear interpolation, and regression-based imputation [2, 3].

Mean imputation replaces missing values with the mean value of the observed data for the corresponding time points. This approach assumes that the missing values are randomly absent and that the mean serves as a representative estimation. Forward/backward filling propagates the last observed value forward or backward to fill in the missing values, assuming that the subsequent or preceding values are similar. Linear interpolation estimates the missing values based on a linear relationship between the neighboring observed values. Regression-based imputation utilizes regression models to predict missing values based on other variables or past observations.

Earlier studies [4] on missing data imputation conducted through bibliometric analysis revealed a growing trend in the field, with a focus on computer science, mathematics, and medical research. The study identified random forest and KNN algorithms as promising techniques for missing data imputation, suggesting the need for future research to compare and evaluate their performance. However, limitations include the exclusive use of the Scopus database and the potential for expanding the analysis to other databases and document types in future studies. This article [5] reviews and analyzes 191 related articles published between 2010 and August 2021 to investigate missing value imputation (MVI) methods for incomplete datasets. The findings reveal that statistical methods such as EMMVI, MMVI, LLSMVI, BPCAMVI, and LRMVI are preferred due to their efficiency. RMSE, NRMSE, MSE, MAE, and R^2 are commonly used evaluation metrics, and KNN, RF, SVM, BPCA, and DT are popular ML models for indirect MVI evaluation. These findings provide valuable insights for researchers in selecting suitable MVI methods and evaluation metrics in real-life applications. It should be emphasized that instead of creating new algorithms, exploring variations of existing imputation approaches is a valuable contribution to the systematic evaluation of current methods. It is important to optimize the execution of algorithms by considering suitable parameters and operating platforms, avoiding the unnecessary development of new imputation techniques before fully exploring the capabilities of existing ones [6].

However, it is important to note that imputation methods can introduce biases and distort the underlying patterns in the time series if the assumptions do not hold or if the imputation technique is not appropriate for the specific dataset. The choice of imputation method should consider the nature of the data, the missing data mechanism, and the specific research or analysis objectives. Improper handling of missing data can greatly affect the accuracy of data-driven insights. This can lead to a reduced sample size and introduce bias, potentially limiting the conclusive findings of the study [7]. Also, combining data from wireless sensor networks (WSNs) can be applied to compare different weight designs of the consensus method and introducing a fully-distributed stopping criterion [8].

In this paper, we present a comparative study of missing value imputation algorithms for time series datasets. We focus on the KNN imputer, the Expectation-Maximization (EM) algorithm, and the Multiple Imputation by Chained Equations (MICE) method. Our study aims to evaluate the performance of these methods in terms of imputation accuracy, considering RMSE and MAE as evaluation metrics. By analyzing and comparing the results, we aim to identify the most effective imputation algorithm for time series datasets, providing valuable insights for researchers and practitioners working with missing values in time series data.

The structure of the paper is as follows: Sect. 2 introduces the definitions of missing value imputation methods that were mentioned previously. The evaluation and comparison of the imputation methods are presented in Sect. 3, along with the experimental results. Finally, Sect. 4 concludes the paper and highlights future research directions.

2 Imputation Methods for Time Series

2.1 K-Nearest Neighbor Imputation

A common imputation technique used to complete missing values in time series datasets is K-Nearest Neighbor (KNN) imputation. Its foundation is the notion that similar data points frequently have comparable values. In KNN imputation, the missing values are estimated by taking into account the features of their closest neighbors. Prior to using top KNN, Dubey A., and Rasool A. [9] employed a few local imputation techniques for bioinformatics' microarray gene expression data. In various domains, researchers have developed improved and modified versions of KNN, such as the CVBkNNI proposed by Huang et al. [10]. This adaptive KNN-based imputation approach selects optimal estimators for missing values and demonstrates superior performance compared to other methods, emphasizing the significance of feature relevance and the influence of missingness mechanisms on imputation accuracy. The study [11] introduces CM-KNN, an improved KNN method for classification, regression, and missing data imputation. It incorporates individualized k values based on prior knowledge and demonstrates robustness to noisy datasets.

The following is how the KNN imputation method operates: To gauge how similar the data points are to one another, a distance metric (such as the Euclidean distance) is first used. The K closest neighbors with the most similar observed values are then found for each missing value. By averaging or weighting the values of these neighbors, the missing value is imputed (see Fig. 1).

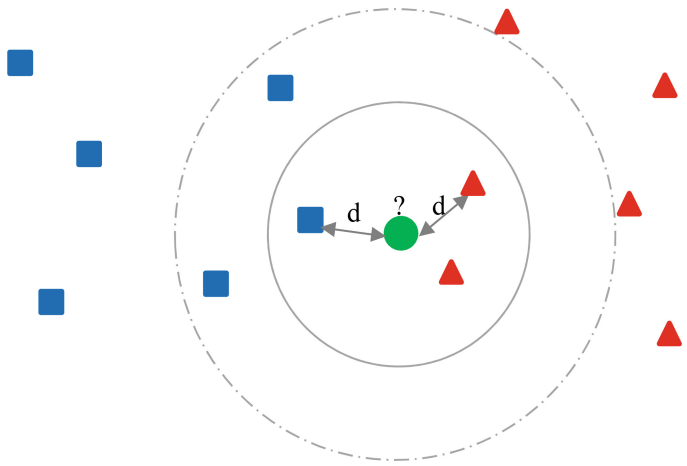


Fig. 1. Missing value imputation using the KNN method with a fixed k value

KNN imputation has a number of benefits. It is a non-parametric method, which means it does not assume anything about how the data are distributed. It is appropriate for a variety of time series data types since it can handle both numerical and categorical variables. Additionally, because KNN imputation takes into account nearby time series points, the temporal structure of the data is preserved [12].

As a result, KNN imputation offers a versatile and efficient imputation strategy for addressing missing values in time series datasets. It is suitable for maintaining the temporal patterns and properties of the time series data since it can take advantage of the similarity between data points.

2.2 Expectation-Maximization Imputation

The EM algorithm operates by alternately computing the expected values of the missing data (the "E-step") and maximizing the likelihood function based on the observed and imputed data (the "M-step"). In the E-step, the algorithm estimates the missing values by calculating the conditional expectations given the observed data and the current parameter estimates. In the M-step, it updates the parameter estimates by maximizing the likelihood function based on the complete dataset, including the imputed values [13, 14].

One of the key advantages of the EM imputation method is its ability to handle complex multivariate data. It is applicable to various types of data, including continuous, categorical, and mixed-type variables. The algorithm can be tailored to different statistical models, such as linear regression, generalized linear models, and multilevel models, allowing researchers to incorporate domain-specific knowledge and assumptions [15].

2.3 Multiple Imputation by Chained Equations

A popular imputation technique called Multiple Imputation by Chained Equations (MICE) deals with the problem of missing data by iteratively predicting the missing values based on the observed data [16]. Regression models are used in the multivariate imputation method (MICE) to impute missing values for various variables. In an iterative procedure, the variables are cycled through repeatedly while the imputed values are updated based on the most recent imputations of other variables. This makes it possible to capture the intricate relationships between variables and to offer imputations that are believable and maintain the data structure.

MICE has gained popularity due to its flexibility in handling missing data across various domains and its ability to account for the uncertainty introduced by imputation. The method is implemented through packages like 'mice' in R and provides a comprehensive framework for imputing missing values in both continuous and categorical variables. The flexibility of MICE allows researchers to incorporate different types of models, such as linear regression, logistic regression, or predictive models, depending on the nature of the data and the research question at hand [17].

3 Experimental Results

3.1 Dataset Description

To illustrate the practical implementation of Missing Value Imputation (MVI) techniques, we present a real-world case of filling gaps in sensor data. The dataset used in this study was collected from the High-Pressure Grinding Rolls (HPGR) in an ore factory located in Iran. The data collection period spanned from June to December 2021.

The dataset comprises primary data obtained from 50 sensors installed on various equipment within the HPGR department. These sensors continuously measure various parameters at 15-min intervals. A specific set of parameters, listed in Table 1, was selected for analysis. The experimental dataset consists of 50 variables in addition to the timestamp.

Table 1. Sensor information regarding to equipment category

Signal Type	Category related to Equipment
Temp	Cooling Tower, HPGR
Pressure transmitter	Compressor, Pump, Cooling Tower, HPGR, Thickener
current	Pump
flow transmitter	HPGR, Flocculant
level transmitter	Mixer, Thickener, HPGR, Thickener, Flocculant
vibration	HPGR
speed	HPGR, Vibratory Feeder

Due to occasional sensor malfunctions or power losses, irregular gaps in the time series data occurred. During the data preprocessing stage, variables with insufficient values for prediction were removed, leaving only the parameters deemed suitable for imputation methods.

3.2 Methodology

In this study, we employed the methodology depicted in Fig. 2 to empirically evaluate the influence of missing value imputation on time series forecasting. The initial stage encompassed data preprocessing procedures, including data cleansing, to generate a refined dataset. To address the missing values within the dataset, we applied various imputation techniques such as KNN, EM, and MICE. Each imputed dataset was then used to train individual models again by applying random missing values. Subsequently, we employed loss functions to evaluate and compare the performance of these forecasting models.

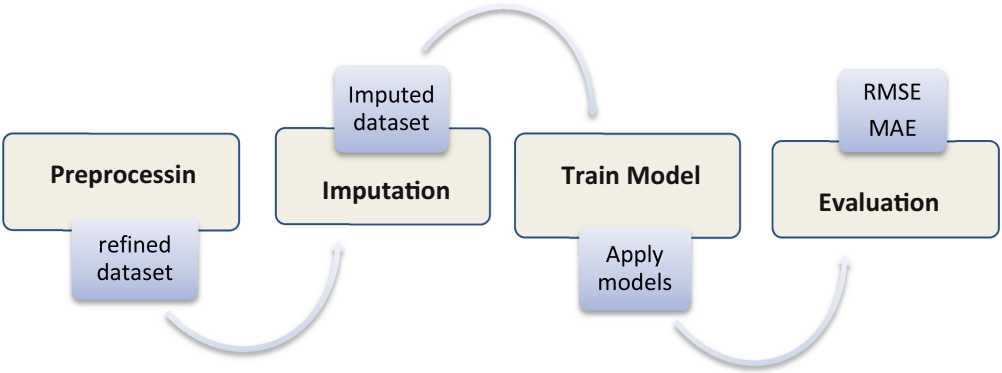


Fig. 2. Methodology procedure

3.2.1 Data Cleaning and Preprocessing

Data cleaning and preprocessing are essential steps in preparing the sensor data for subsequent analysis and training. In this section, we describe the procedures used to handle missing values and outliers in the dataset.

Missing Value Handling:

To address missing values, we first examine the dataset to identify the presence of NaNs (missing values). Figure 3 displays the distribution of NaNs across the dataset, obtained by counting the number of NaNs in each column using the isna() and sum() methods. We set a threshold for the minimum number of non-null values (e.g., thresh = len(data) * 0.5) to retain columns with more than 50% non-null values. The dropna() method is then employed with the specified threshold and axis = 1 parameters to remove columns exceeding the threshold. The resulting Data Frame, denoted as data_clean, contains only the columns with less than or equal to 50% NaNs.

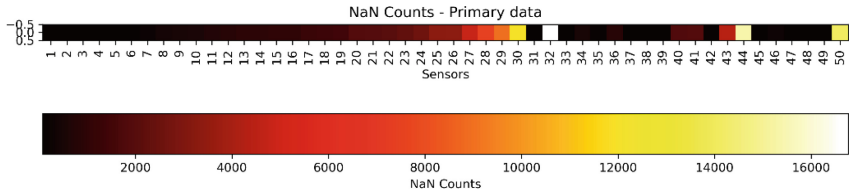


Fig. 3. Distribution of NaNs across the dataset

After dropping the columns (see Fig. 4), we evaluate the impact of this operation by calculating the percentage of missing values before and after the cleaning process. The initial percentage of missing values is shown as 13.92, and after dropping the columns, the percentage is reduced to 6.76.

Furthermore, to gain insights into the missing pattern, we utilize a heatmap visualization (see Fig. 5). A heatmap of missing values represents the distribution and location of missing values in the dataset. Each cell in the heatmap corresponds to a value in

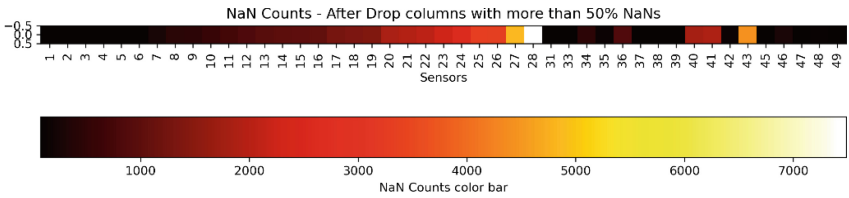


Fig. 4. Distribution of NaNs across the dataset after drop columns

the dataset, with white cells indicating missing values and colored cells denoting non-missing values. The visualization helps identify the extent and distribution of missing values, aiding in the selection of appropriate imputation methods. The left side of the figure depicts the NaN values before cleaning, while the right side shows the heatmap of missing values after cleaning. The horizontal bar denotes the sensor IDs.

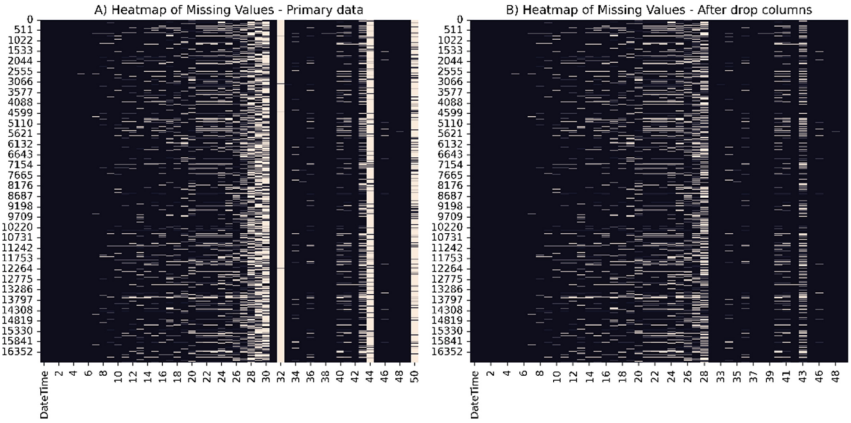


Fig. 5. A heatmap comparison of both primary data (A) and after removing columns data (B)

Robust Outlier Handling:

Another crucial step in data preprocessing is outlier removal. In this study, we employed a different approach to handle outliers compared to the previous method. Instead of removing entire rows with outliers, we utilized the Interquartile Range (IQR) method to replace the outlier values with the corresponding lower or upper whisker values.

The IQR method calculates the range between the first quartile (Q1) and the third quartile (Q3) of the data distribution. Any value below Q1 - 1.5 * IQR or above Q3 + 1.5 * IQR is considered an outlier. Rather than discarding the entire row, we replace these outlier values with the nearest lower or upper whisker value.

This approach allows us to retain more data compared to the previous method, where rows with outliers were completely removed. By replacing outliers with the respective whisker values, we aim to mitigate the impact of outliers while preserving a larger portion of the dataset for further analysis. As an example, we present Fig. 6 representing the data

of a specific sensor before and after refining data quality through robust outlier handling. The figure clearly demonstrates the impact of the refining process on the sensor data.

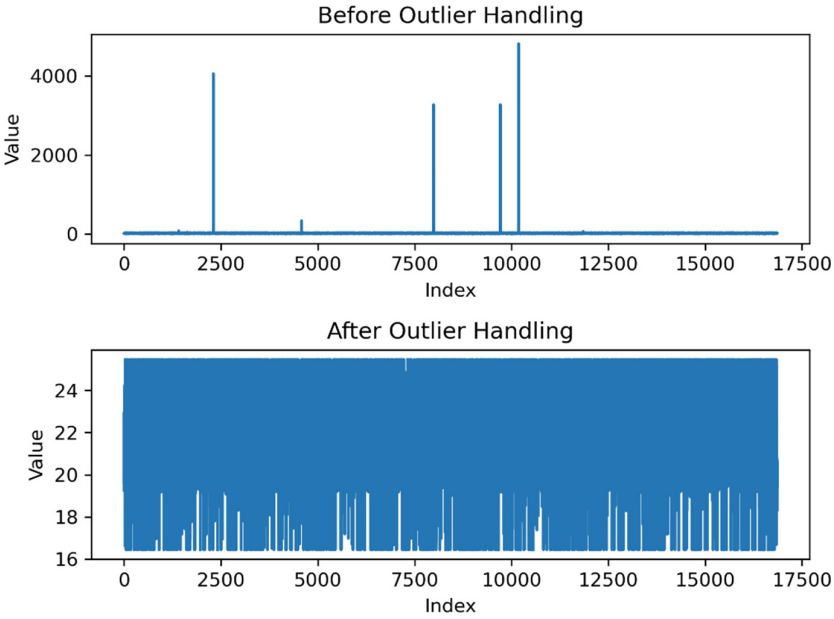


Fig. 6. Comparison of dataset before and after Handling Outliers for one specific sensor

3.2.2 Imputation Framework

In order to address missing values in the dataset and enable accurate analysis, we employed a comprehensive imputation framework that encompasses three different methods: KNN, EM, and MICE. Each method provides a unique approach to impute missing values, catering to different characteristics of the data.

For the KNN method, we utilized the KNN imputer. By considering the nearest neighbors, the missing values were imputed based on the values of the neighboring data points. We set the number of neighbors to 5 and applied the imputation process to all columns except the first one. Similarly, the EM (Expectation-Maximization) method was implemented using the Iterative. This method iteratively estimates the missing values based on the observed values and maximizes the likelihood of the data. The last method we utilized was Multiple Imputation by Chained Equations (MICE). MICE is a powerful imputation technique for time series data that generates multiple imputed datasets by iteratively modeling the missing values based on the observed data. In our study, we implemented the MICE algorithm to impute the missing values in each column. This approach allows us to capture the dependencies and relationships within the time series data while imputing the missing values. By iteratively updating the imputed values based on the imputed values from other variables, MICE provides a comprehensive imputation solution for time series datasets.

Figure 7 presents the imputation methods applied to a specific sensor within the dataset. The figure showcases the comparison between three different imputation techniques, including refined data: Refined Data Before Imputation (A), KNN (B), EM (C), and MICE (D). Subplots B, C, and D in Fig. 7 represent the data after applying a specific imputation method to fill in the missing values. However, due to the varying amount of data and the distinct range of sensors, it can be challenging to visually discern the differences between the imputation techniques. Therefore, in the subsequent section, the results will be presented based on the calculation of RMSE and MAE metrics. These quantitative measures will provide a more comprehensive evaluation of the performance of each imputation method, enabling a more objective assessment of their effectiveness in handling missing data for the given sensor.

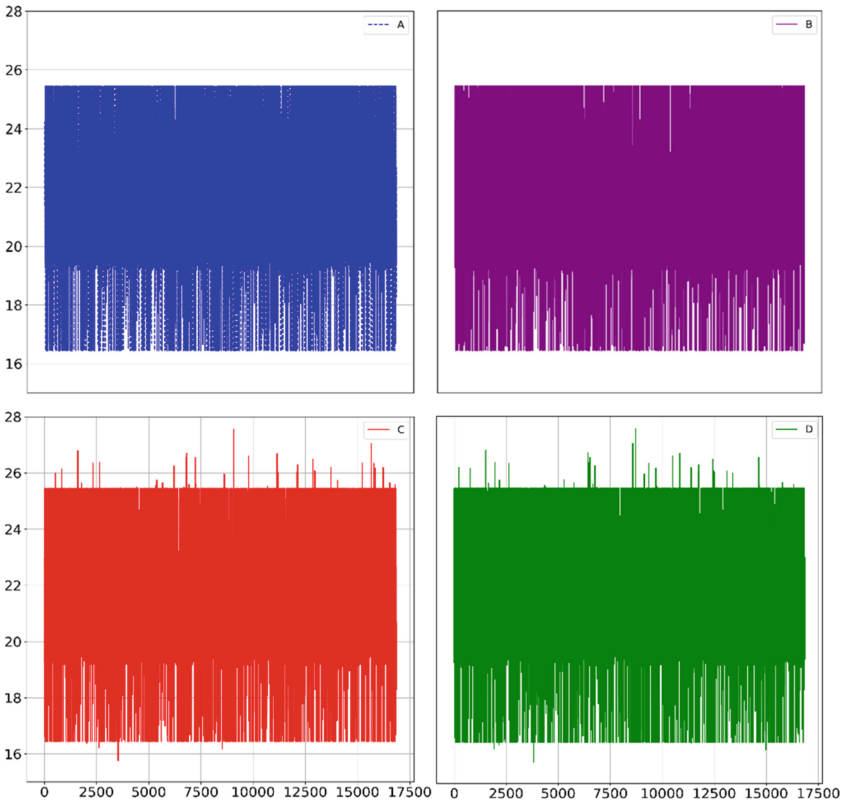


Fig. 7. Comparison of Imputation Methods with Refined Data on the specific sensor: Refined Data Before Imputation (A) vs. KNN (B) vs. EM (C) vs. MICE (D)

As our aim was to compare the performance of these methods and identify the most suitable approach for imputing missing data to evaluate the imputation performance, we introduced random missing values into the imputed datasets again. The missing values were randomly distributed across the data, accounting for a specified percentage of the

total values. This process was essential to simulate realistic scenarios with missing data, as missingness is often non-random and can impact subsequent analyses.

For each imputation method, we created copies of the original imputed datasets and applied the respective imputation technique to fill in the missing values with the same parameters one more time. After imputing the missing values, we compared the imputed datasets with the corresponding original datasets to assess the quality of the imputation.

In order to provide a comprehensive comparison of the original imputed, and new imputed datasets using the KNN, EM, and MICE methods, we visualized the data using box plots. Figure 8 displays the box plots for a specific sensor in the respective datasets.

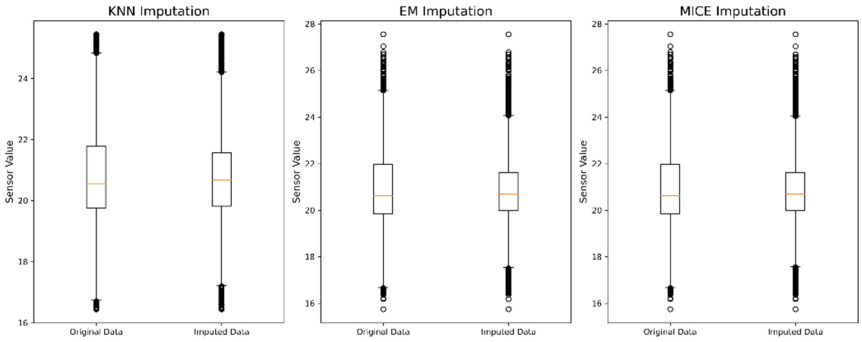


Fig. 8. Comparing Final Imputation Results: Original Imputed vs. Newly Imputed Datasets

In the subsequent section, the final outcomes will be analyzed and evaluated using RMSE and MAE metrics, aiming to determine the optimal imputation approach.

3.3 Results and Evaluation

Evaluation metrics such as Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) are commonly utilized to assess the performance of imputation methods. RMSE measures the average difference between the predicted and actual values, providing an indication of the overall accuracy of the imputed data. A lower RMSE value indicates a better fit between the imputed and actual values. On the other hand, MAE quantifies the average absolute difference between the predicted and actual values, highlighting the average magnitude of errors. Similarly, a lower MAE value suggests a more accurate imputation. They are defined as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2} \tag{1}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right| \tag{2}$$

where n is the number of observations, y_i is the observed values, and $\hat{y}_i$ is the estimated missing value. A small value as an output for these performance metrics means that the estimated value is close to the real value.

In our study, we employ RMSE and MAE as evaluation metrics to compare the performance of different imputation methods (for all 45 sensors), namely KNN, EM, and MICE. These methods were applied to impute missing values in the original datasets, resulting in new imputed datasets. Subsequently, the imputed datasets were evaluated using RMSE and MAE to determine the effectiveness of each imputation approach. The evaluation results presented in Table 2 contribute to the selection of the most suitable imputation method.

Table 2. RMSE and MAE Evaluation Results for Imputation Methods for All Sensors

Method	RMSE	MAE
KNN	8.495	1.003
EM	8.559	1.008
MICE	8.493	1.002

By analyzing the RMSE and MAE values, we can determine the imputation method that demonstrates superior performance in handling missing data. In this case, based on the obtained results, we observe that both KNN and MICE methods exhibit similar performance, with slightly lower RMSE and MAE values (KNN – RMSE: 8.495 MAE: 1003, MICE – RMSE: 8.493 MAE: 1002) compared to the EM method (RMSE: 8.559 MAE: 1008). Therefore, KNN and MICE methods can be considered more suitable options for imputing missing values in our dataset.

4 Conclusion

The evaluation of imputation methods for handling missing values in time series forecasting is a crucial aspect of data analysis [3]. The data imputation methods suit well the scenarios for real-time data collection [18], when the sensor device or collection system is under maintenance or has atemporal failure. In this study, we adopted a novel approach to evaluate these methods by utilizing the original dataset for 50 sensors with real missing values. We then applied imputation techniques based on the existing missing value patterns and compared the newly imputed datasets with the original imputed dataset to determine the most effective method.

The results revealed that MICE and KNN outperformed the other method (EM) in terms of model performance. However, it is important to acknowledge the limitations of these findings. To further enhance the evaluation process, future research should incorporate a more comprehensive benchmark study that encompasses a wider range of imputation approaches, including machine learning techniques and considers various missing-data scenarios. Additionally, conducting experiments to examine the effectiveness of the imputation process in reconstructing missing values introduced through simulation will provide a more comprehensive understanding of the imputation effects.

Acknowledgments. This work was supported by the Slovak Scientific Grand Agency VEGA under the contract 2/0135/23 “Intelligent sensor systems and data processing” and “Research on

the application of artificial intelligence tools in the analysis and classification of hyperspectral sensing data" (ITMS: NFP313011BWC9) supported by the Operational Programme Integrated Infrastructure (OPII) funded by the ERDF.

References

1. Little, R.J., Rubin, D.B.: Statistical Analysis with Missing Data. Wiley, Hoboken (2019)
2. Emmanuel, T., Maupong, T., Mpoeleng, D., Semong, T., Mphago, B., Tabona, O.: A survey on missing data in machine learning. J. Big Data. **8** (2021). https://doi.org/10.1186/s40537-021-00516-9
3. Ahn, H., Sun, K., Kim, K.P.: Comparison of missing data imputation methods in time series forecasting. Comput. Mater. Continua. **70**, 767–779 (2021). https://doi.org/10.32604/cmc.2022.019369
4. Jamaludin, K.R., Muhamad, W.Z.A.W., Miskon, S.: A review of current publications trend on missing data imputation over three decades: direction and future research (2021). https://doi.org/10.21203/rs.3.rs-996596/v1
5. Hasan, M.K., Alam, M.A., Roy, S., Dutta, A., Jawad, M.T., Das, S.: Missing value imputation affects the performance of machine learning: a review and analysis of the literature (2010–2021) (2021). https://doi.org/10.1016/j.imu.2021.100799
6. Armina, R., Mohd Zain, A., Ali, N.A., Sallehuddin, R.: A review on missing value estimation using imputation algorithm. J. Phys. Conf. Ser. (2017). https://doi.org/10.1088/1742-6596/892/1/012004
7. Read, S., Wild, S., Lewis, S.: Applying missing data methods to routine data: a prospective, population-based register of people with diabetes. Trials. **14** (2013). https://doi.org/10.1186/1745-6215-14-s1-p113
8. Kenyeres, M., Kenyeres, J.: Multi-sensor data fusion by average consensus algorithm with fully-distributed stopping criterion: comparative study of weight designs. U.P.B. Sci. Bull., Series C. **81** (2019)
9. Dubey, A., Rasool, A.: Efficient technique of microarray missing data imputation using clustering and weighted nearest neighbour. Sci. Rep. 11, (2021). https://doi.org/10.1038/s41598-021-03438-x
10. Huang, J., et al.: Cross-validation based K nearest neighbor imputation for software quality datasets: an empirical study. J. Syst. Softw. **132**, 226–252 (2017). https://doi.org/10.1016/j.jss.2017.07.012
11. Zhang, S., Li, X., Zong, M., Zhu, X., Cheng, D.: Learning k for kNN Classification. ACM Trans. Intell. Syst. Technol. **8** (2017). https://doi.org/10.1145/2990508
12. Zhang, S.: Parimputation: from imputation and null-imputation to partially imputation. IEEE Intell. Inf. Bull. **9**, 32–38 (2008)
13. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the EM algorithm. (1977)
14. Nakai, M., Ke, W.: Review of the methods for handling missing data in longitudinal data analysis (2011)
15. Latiffah Abd Rani, N., et al.: Prediction model of missing data: a case study of PM10 across Malaysia Region. Article J. Appl. Fundam. Sci. **2018**, 182–203 (2019). https://doi.org/10.4314/jfas.v10i1s.1
16. van Buuren, S.: Flexible Imputation of Missing Data. Chapman and Hall/CRC (2012). https://doi.org/10.1201/b11826

17. Azur, M.J., Stuart, E.A., Frangakis, C., Leaf, P.J.: Multiple imputation by chained equations: what is it and how does it work? Int. J. Methods Psychiatr. Res. **20**, 40–49 (2011). https://doi.org/10.1002/mpr.329
18. Gatial, E., Balogh, Z., Hluchy, L.: Concept of energy efficient ESP32 chip for industrial wireless sensor network. In: 2020 IEEE 24th International Conference on Intelligent Engineering Systems (INES), pp. 179–184. IEEE (2020). https://doi.org/10.1109/INES49302.2020.9147189

Motivation to Learn in an E-learning Environment with Fading Mark

Roman Tsarev[1,2(✉)], Younes El Amrani[3], Shadia Hamoud Alshahrani[4], Naim Mahmoud Al Momani[5], Joel Ascencio[6], Aleksey Losev[7], and Kirill Zhigalov[1,8]

[1] MIREA - Russian Technological University (RTU MIREA), Moscow, Russia
tsarev.sfu@mail.ru
[2] Bauman Moscow State Technical University, Moscow, Russia
[3] Abdelmalek Essaâdi University, Tangier, Morocco
[4] Medical Surgical Nursing Department, King Khalid University, Khamis Mushate, Saudi Arabia
[5] Al Ain University, Abu Dhabi, UAE
[6] Universidad César Vallejo Lima Perú, Los Olivo, Peru
[7] Russian State Agrarian University - Moscow Timiryazev Agricultural Academy, Moscow, Russia
[8] V.A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia

Abstract. E-learning is an innovative trend in education today, information and communication technologies have significantly changed the educational process. The use of electronic educational environment, as the main tool of e-learning, has a number of significant advantages associated with the organization of learning, monitoring and evaluation of educational results. The high level of the required final competences of students depends on the degree of their motivation, internal and external motives, self-discipline, self-organization and self-control. There is a tendency of inefficient time allocation among students of higher education institutions, that is why we have developed a special method of assessment, thanks to which students will be motivated to pass academic work on time and get high marks. #COMESYSO1120.

Keywords: e-learning environment · e-learning platform · motivation · grading system

1 Introduction

Currently, there is a rapid development of e-learning, which is actively integrated into the modern system of education [1, 2]. Learning today is no longer imaginable without the use of information and communication technologies, additional digital resources and interactive methods [3–5]. E-Learning environment ensure distance interaction between the participants of the educational process, contribute to individualization of learning and selection of individual educational trajectory, provide access to the curricula, electronic resources and the results of midterm assessment [6–9].

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 91–99, 2024.
https://doi.org/10.1007/978-3-031-53552-9_8

The most important structural component of the e-learning environment is educational platforms [8–11]. The main didactic features of electronic educational platforms include the use of modern information and communication technologies and multimedia, which provide visibility and accessibility of educational material; the ability to develop self-organization, discipline and initiative; comprehensive consideration of educational material; interaction with the teacher regardless of place and time; the possibility of practical application of the knowledge obtained; implementation of the diagnostic function [12]. The last didactic feature optimizes the process of control and assessment of the obtained knowledge of students due to automatic verification of test tasks and the exclusion of the subjective factor in the verification [13].

Higher education institutions use electronic educational platforms of their own design or ready-made solutions, such as LMS Moodle [14, 15]. Such modern software allows a teacher and students to effectively interact online. Such platforms create electronic educational courses for distance and blended learning, develop modules with interactive material, assignments and tests [16–19]. The concept of Moodle is not associated with the denial of traditional forms and methods of learning, it implements two formats - blended and distance learning.

The obvious advantage of electronic educational media is the absence of the need to be in the classroom of the university for the course [20]. However, this advantage turns out to be a disadvantage, because the factor of external motivation and discipline, which is typical for traditional classes in the classroom, disappears. Students, getting more freedom in learning, often are not able to intelligently organize their time. The lack of self-discipline also becomes a problem, which manifests itself by the time of exams or tests. It is important for the instructor to organize the work in such a way that the students understand, so that the student understands his/her responsibility in relation to the norms and rules of education [21–24].

Academic success (a student's ability to achieve their goals, to realize themselves, to get satisfaction from their actions) directly depends on personal characteristics: motivation, self-discipline, self-organization, self-control [25–28].

Motivation for learning is one of the most significant non-cognitive factors in achieving high results and is a complex of internal and external motives [26, 29, 30]. Many researchers believe that high learning motivation can exceed even the contribution of intelligence in achieving academic success. The source of intrinsic motivation is the learner himself, it arises through the arisen pleasure and interest in cognitive activity, its result. The desire for self-development or development of certain personal qualities and abilities, the need for intellectual activity and overcoming learning difficulties also belong to internal motives. External motives are associated with aspirations that stimulate to perform learning tasks for the sake of reward, encouragement, praise, prestige, receiving high marks. Duty and responsibility are also extrinsic motivations for learning. Thus, extrinsic motivation is closely related to the assessment of student performance [31–33].

Assessment implies the implementation of several types of control: current, midterm and final attestation. Electronic testing, as a result of studying the educational platform course, is a convenient intersubjective way of evaluating results, it provides an independent assessment of student learning achievements to determine the level of knowledge

of students; evaluation of the effectiveness and efficiency of the educational process and the use of electronic educational environment; continuous monitoring of student knowledge.

Traditional assessment system does not always reflect the real situation of the skill developed by students, does not form evaluative independence of students. Therefore, in pedagogical practice teachers strive to improve the means and forms of control, to move to a point system of evaluation in order to increase objectivity and reliability of assessment. Many researchers have proved that the use of point grading system increases the level of motivation of students to learn. The point-based grading system involves evaluating the results in points, the number of which is set by the teacher. These points can be gained by the student during lectures, practical or laboratory classes. It is important to note that with this grading system the student knows in advance what number of points he or she will receive for a particular type of work in case of successful completion of assignments. Other advantages of the grading system include:

- activation of cognitive activity of students;
- stimulation for better performance of tasks during the whole semester;
- individualization of the learning process;
- formation of rational approach to learning;
- increasing objectivity in the evaluation of knowledge.

Assessment within the electronic educational environment takes place in accordance with the passage of modules, i.e. the study of each module ends with intermediate control in the form of different types of assignments. Thus, the student accumulates points, which are then added up and based on the results of the resulting sum of points a credit or examination grade is given. Therefore, such a system is called modular-rating and allows you to assess the student's workload in the process of mastering the course.

Sometimes the teacher can provide work for extra points, including creative assignments, such an opportunity to adjust their learning activities encourages students to learn and to show creative abilities.

The method of rating carried out with the help of a point grading system is also a stimulating incentive and is often used in e-learning environments. Students can correlate their results with the results of their classmates, high scores (lines in the rating) raise the students' authority, increase their self-esteem, and motivate them to achieve high results.

So, the formation of internal and external motives is an essential issue for a teacher, but it is worth noting that the main motives of university students are external reasons - control of class attendance, the threat of sanctions from the teacher and the dean's office, passing a test or exam to get a diploma, the possibility of exclusion from the educational institution. The desire to avoid negative evaluations is a driver of learning. Motivation and academic success in general are closely related to students' self-discipline and conscientiousness.

In the process of learning with the help of electronic educational environment the role of students' independent activity increases. Independent work is a form of joint unified activity of a teacher and students. Performing independent work, students actively operate the acquired knowledge, skills and abilities, perform exploratory activities. Therefore, independent activity of students, carried out with the help of electronic educational environment, forms the consciousness of students.

Self-discipline implies the ability to manage one's desires, refrain from impulsive behavior. Self-organization is determined by the ability to independently organize their learning activities, effectively manage their time (the ability to distribute it properly), and use the available resources wisely. Electronic educational platforms allow self-control of learners: to set a goal and choose methods of achieving the expected result, to determine their weaknesses and causes of failures, to correct their independent learning activities, to analyze and evaluate the final result.

Thus, within the framework of e-learning, there is a need for some external factor that would promote students' motivation to follow the program of the course in a timely manner, encourage students to study theoretical material and pass practical assignments in time, assumed by the plan of study of the course.

2 Method

The paper proposes a method of assessing the progress of a student passing an electronic course, which is based on such criteria as the quality of assimilation of material and, mainly, and the timeliness of passing of completed practical assignments. An electronic educational course is divided into modules according to the calendar and thematic plan of the discipline (a semester consists of 16–18 weeks, depending on the university). If there are 16 weeks, with the frequency of 1 practical lesson every week, we have 8 practical lessons. Under the methodology presented, it is assumed that the mark will decrease (the farther away, the more) as the expected time of passing of the assignment is distanced. Completion of all assignments in time within the module allows the student to gain 100 points, as a maximum mark. The 100 points grading system, in contrast to the traditional 5-point system, is more effective in obtaining an objective overview of the obtained knowledge of students, it is convenient and understandable, is a means of additional motivation for the study of the course. In addition, such a mark can easily be reduced to any other grading system.

So, the period when a student needs to pass the assignment, denote it t_0, we have a mark set by the teacher, multiplied by the coefficient of 1.0. The next period, for example the next two weeks, let's denote it by t_1. We multiply the mark obtained by 0.9, i.e. it is decreased by 10%. At the next period t_2 we multiply the obtained score by 0.8. At each subsequent time interval the coefficient is decreased until it reaches the value of 0.6.

It is assumed that the need to save points and obtain a high mark will motivate students to allocate time efficiently and, as a consequence, to pass the assignment on time. Note that for the teacher facilitates the process of evaluating the results of students, there is no need to calculate the mark with the deadline for passing the report on the practical assignment due to the automation of the process.

During this study, the evaluation system in LMS Moodle was configured so as to recalculate the final mark, based on the evaluation set by the teacher and the time of delivery of the report on practical work.

The assessment criteria included the quality of the report provided, the depth of understanding of the studied material, the student's ability to answer additional questions, the design of the work, etc. Assessment in points, which is put by the teacher, can be

expressed as follows:

$$m = c_1 \cdot x_1 + c_2 \cdot x_2 + ... + c_n \cdot x_n,$$

where x_i, $i = 1..n$ is the factor related to the practical assignment being assessed by the teacher; c_i, $i = 1..n$ is the corresponding normalized coefficient.

The final mark for the practical assignment is based on the teacher's assessment of the assignment and the deadline for passing the report on the practical assignment. It is calculated as follows:

$$M = \begin{cases} m \cdot 1.0 \text{ for } t_0 \\ m \cdot 0.9 \text{ for } t_1 \\ m \cdot 0.8 \text{ for } t_2 \\ m \cdot 0.7 \text{ for } t_3 \\ m \cdot 0.6 \text{ for } t_4 \end{cases} \quad (1)$$

where M is the final mark for the assignment; m is the mark provided by the teacher; t_0,…, t_4 shows the term of passing the report on the practical assignment, t_0 corresponds to timely passing, the higher the index of the variable t, the greater the delay in passing the report. Introducing a coefficient lower then 0.6 is impractical, because this will not allow the student to receive even a minimum mark for the practical assignment.

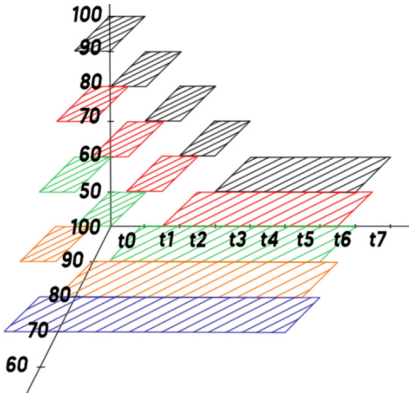


Fig. 1. Assessment of the practical assignment of the student taking into account the term of its passing.

Visually, the grading system, taking into account the time of passing the report on the practical assignment is shown in Fig. 1.

Further, if there is a need, the marks received during the semester can be converted to another grading system: 5, 4, 3, 2; A, B, C, D, F; "excellent", "good", "satisfactory" or any other.

3 Results and Discussion

The application of the proposed grading system with motivation for timely passing of reports on practical assignments allowed to significantly increase the involvement of students in the educational process, improve their organization and motivate them to pass practical assignments according to the plan of study of the discipline.

Only 4% of 132 students for one reason or another did not adhere to the expected deadlines for studying certain topics of the course and passing of practical assignments on time. 9% of students were late for one period (t_1) and 3% of the students were late for two periods (t_2). There were no students who were late for three or four periods (t_3 and t_4).

So, it can be stated that the proposed approach significantly increases students' motivation to follow the course study plan and increases the effectiveness of learning the course.

The experiment took place within the traditional form of learning, but with the use of electronic learning environment. We assume that this methodology will also be very effective in distance learning, in which the factor of the presence of the teacher in the classroom or the factor of the presence of the student in the class at a certain time are absent.

It should also be noted the flexibility of LMS Moodle in setting the assessment of the student in accordance with the proposed methodology. This system allowed making the required changes, which made it possible to take into account the deadlines for the practical assignment, and also relieved the teacher of the need to make this recalculation manually.

4 Conclusion

Development of e-learning environment is one of the priorities of the educational activity of each university. Any innovations in the field of e-learning should have a solid foundation, which consists of effective and interactive methods and means of learning. These include electronic educational environments and educational platforms. They provide independence of learning from place and time, access to curricula, programs of disciplines and lectures, automated control of learning outcomes. Assessment is a quality control of education, a way to correlate the learning activities of students, a means to determine the development and progress of teaching activities. It is the stimulating function of assessment that motivates students, encourages, inspires confidence in the achievability of new goals, obtaining a higher level of knowledge.

The issue of student learning motivation within the use of e-learning environment is a burning issue of higher education. Most students, for one reason or another, do not know how to effectively manage their time, are not interested in performing learning assignments, and are passive in the course of studies. Thus, we can say about the academic negligence of modern students, which reduces the quality and effectiveness of education in general. Of course, there are also such external factors as stress, examination tension, lack of time, affecting learning and grades.

Proper organization of the learning process, structuring of learning material and encouragement can increase the level of motivation of students. It is important for the

instructor to help students organize and allocate their time, develop an incentive strategy in which all work will be handed in on time. Therefore, to increase students' motivation we developed our own grading methodology, the basis of which was the assumption that lowering the mark as the distance from the expected moment of passing the assignment will motivate students to complete practical assignments on time. According to the results of the research the hypothesis was confirmed, a significant majority of students were able to complete the assignment on time, which confirms the greater involvement of students in the educational process and their motivation both to study the course and to pass timely reports on completed practical assignments.

References

1. Baig, M.I., Shuib, L., Yadegaridehkordi, E.: E-learning adoption in higher education: a review. Inf. Dev. **38**(4), 570–588 (2022). https://doi.org/10.1177/02666669211008224
2. Gupta, A., Motwani, S., Agarwal, A., Udandarao, V., Chakraborty, T.: Changing landscape of technical education pedagogy from traditional to practical e-learning. Computer **55**(11), 16–28 (2022). https://doi.org/10.1109/MC.2022.3164231
3. Yurchenko, P.: Ways to solve the problem of documentary thematic search. Inf. Econ. Manag. **2**(1), 0101–0123 (2023). https://doi.org/10.47813/2782-5280-2023-2-1-0101-0123
4. Lunev, D., Poletykin, S., Kudryavtsev, D.O.: Brain-computer interfaces: technology overview and modern solutions. Modern Innov. Syst. Technol. **2**(3), 0117–0126 (2022). https://doi.org/10.47813/2782-2818-2022-2-3-0117-0126
5. Zenyutkin, N., Kovalev, D., Tuev, E., Tueva, E.: On the ways of forming information structures for modeling objects, environments and processes. Modern Innov. Syst. Technol. **1**(1), 10–22 (2021). https://doi.org/10.47813/2782-2818-2021-1-1-10-22
6. Deetjen-Ruiz, R., et al.: Applying ant colony optimisation when choosing an individual learning trajectory. In: Silhavy, R., Silhavy, P. (eds.) Networks and Systems in Cybernetics, CSOC 2023, LNNS, vol. 723, pp. 587–594. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-35317-8_53
7. Minamatov, Y E.O.G.L., Nasirdinova, M.H.Q.: Application of ICT in education and teaching technologies. Sci. Prog. **3**(4), 738–740 (2022)
8. Perifanou, M., Economides, A.A.: The landscape of MOOC platforms worldwide. Int. Rev. Res. Open Dist. Learn. **23**(3), 104–133 (2022). https://doi.org/10.19173/irrodl.v23i3.6294
9. Wu, B., Wang, Y.: Formation mechanism of popular courses on MOOC platforms: a configurational approach. Comput. Educ. **191**, 104629 (2022). https://doi.org/10.1016/j.compedu.2022.104629
10. Dong, Y., Shao, B., Lou, B., Ni, C., Wu, X.: Status and development of online education platforms in the post-epidemic era. Procedia Comput. Sci. **202**, 55–60 (2022). https://doi.org/10.1016/j.procs.2022.04.008
11. Veeramanickam, M.R.M., Ramesh, P.: Analysis on quality of learning in e-learning platforms. Adv. Eng. Softw. **172**, 103168 (2022). https://doi.org/10.1016/j.advengsoft.2022.103168
12. Akhmetjanov, M., Ruziev, R.: Fundamentals of modeling fire safety education. Inf. Econ. Manag. **1**(2), 0301–0308 (2022). https://doi.org/10.47813/2782-5280-2022-1-2-0301-0308
13. Tsarev, R., et al.: Improving test quality in e-learning systems. In: Silhavy, R., Silhavy, P. (eds.) Networks and Systems in Cybernetics, CSOC 2023, LNNS, vol. 723, pp 62–68. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-35317-8_6
14. David, A., Mihai, D., Mihailescu, M.-E., Carabas, M., Tapus, N.: Scalability through distributed deployment for moodle learning management system. Procedia Comput. Sci. **214**, 34–41 (2022). https://doi.org/10.1016/j.procs.2022.11.145

15. De Medio, C., Limongelli, C., Sciarrone, F., Temperini, M.: MoodleREC: a recommendation system for creating courses using the moodle e-learning platform. Comput. Hum. Behav. **104**, 106168 (2020). https://doi.org/10.1016/j.chb.2019.106168
16. Gushchin, A.: Algorithmic approach to the design of e-learning courses. In: Silhavy, R. (eds.) Informatics and Cybernetics in Intelligent Systems, CSOC 2021, LNNS, vol. 228, pp. 207–214. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-77448-6_19
17. Kamunya, S.M., Oboko, R.O., Maina, E.M., Miriti, E.K.: A systematic review of gamification within e-learning. In: Handbook of Research on Equity in Computer Science in P-16 Education, pp. 201–218. IGI Global, Hershey, Pennsylvania, US (2020). https://doi.org/10.4018/978-1-7998-4739-7.ch012
18. Poondej, C., Lerdpornkulrat, T.: Gamification in e-learning: a moodle implementation and its effect on student engagement and performance. Interact. Technol. Smart Educ. **17**(1), 56–66 (2019). https://doi.org/10.1108/ITSE-06-2019-0030
19. Tsarev, R., et al.: Gamification of the graph theory course. Finding the shortest path by a greedy algorithm. In: Silhavy, R., Silhavy, P. (eds.) Networks and Systems in Cybernetics, CSOC 2023, LNNS, vol. 723, pp. 209–216. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-35317-8_18
20. Kononenko, A., Kravchenko, M., Nedospasova, L., Fedorovich, E.: E-learning online platforms for educational approach. In: Guda, A. (eds.) Networked Control Systems for Connected and Automated Vehicles, NN 2022, LNNS, vol. 510, pp. 1089–1096. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-11051-1_111
21. Gushchin, A.N., Divakova, M.N.: Nurturing cognitive skills in undergraduates with the help of ontological analysis. IOP Conf. Ser. Mater. Sci. Eng. **972**, 012059 (2020). https://doi.org/10.1088/1757-899X/972/1/012059
22. Leoste, J., et al.: Environment challenges of e-learning in higher education—the teachers' perspective. Smart Innov. Syst. Technol. **908**, 143–156 (2023). https://doi.org/10.1007/978-981-19-5240-1_10
23. Nakajima, T.M., Goode, J.: Transformative learning for computer science teachers: examining how educators learn e-textiles in professional development. Teach. Teach. Educ. **85**, 148–159 (2019). https://doi.org/10.1016/j.tate.2019.05.004
24. Ung, L.-L., Labadin, J., Mohamad, F.S.: Computational thinking for teachers: development of a localised e-learning system. Comput. Educ. **177**, 104379 (2022). https://doi.org/10.1016/j.compedu.2021.104379
25. Gushchin, A.N.: Design-based science: curriculum for architects. AIP Conf. Proc. **2657**, 020047 (2022). https://doi.org/10.1063/5.0107174
26. Huang, A.Y.Q., Lu, O.H.T., Yang, S.J.H.: Effects of artificial intelligence-enabled personalized recommendations on learners' learning engagement, motivation, and outcomes in a flipped classroom. Comput. Educ. **194**, 104684 (2023). https://doi.org/10.1016/j.compedu.2022.104684
27. Pheng, H.S., Chin, T.A., Lai, L.Y., Choon, T.L.: E-Learning as a supplementary tool for enhanced students' satisfaction. AIP Conf. Proc. **2433**, 030005 (2022). https://doi.org/10.1063/5.0072901
28. Rasheed, H.M.W., He, Y., Khalid, J., Khizar, H.M.U., Sharif, S.: The relationship between e-learning and academic performance of students. J. Public Aff. **22**(3), e2492 (2022). https://doi.org/10.1002/pa.2492
29. Tan, L.S., Kubota, K., Tan, J., Kiew, P.L., Okano, T.: Learning first principles theories under digital divide: effects of virtual cooperative approach on the motivation of learning. Educ. Chem. Eng. **40**, 29–36 (2022). https://doi.org/10.1016/j.ece.2022.04.003
30. Jääskä, E., Lehtinen, J., Kujala, J., Kauppila, O.: Game-based learning and students' motivation in project management education. Project Leadersh. Soc. **3**, 100055 (2022). https://doi.org/10.1016/j.plas.2022.100055

31. Alsadoon, E., Alkhawajah, A., Suhaim, A.B.: Effects of a gamified learning environment on students' achievement, motivations, and satisfaction. Heliyon **8**(8), e10249 (2022). https://doi.org/10.1016/j.heliyon.2022.e10249
32. Alyoussef, I.Y.: Acceptance of e-learning in higher education: the role of task-technology fit with the information systems success model. Heliyon **9**(3), e13751 (2023). https://doi.org/10.1016/j.heliyon.2023.e13751
33. Fandiño, F.G.E., Muñoz, L.D., Velandia, A.J.S.: Motivation and e-learning english as a foreign language: a qualitative study. Heliyon **5**(9), e02394 (2019). https://doi.org/10.1016/j.heliyon.2019.e02394

Hub Operation Pricing in the Intermodal Transportation Network

Alexander Krylatov[1,2(✉)] and Anastasiya Raevskaya[1]

[1] Saint Petersburg State University, Saint Petersburg, Russia
a.krylatov@spbu.ru, aykrylatov@yandex.ru
[2] Institute of Transport Problems RAS, Saint Petersburg, Russia

Abstract. The growing importance of strategic supply chain management in light of intermodal transportation networks has attracted increased attention from researchers in recent decades. The study of intermodal logistic services has become urgent for multiple branches of science. This work is focused on hub operation pricing under an equilibrium freight flow assignment model. We formulate this model as a non-linear optimization problem and show that its solution corresponds to the equilibrium assignment pattern in an intermodal transportation network. We analyze the sensitivity of the average purchase cost and hub load to different strategies of hub operation pricing. To this end, we obtain the equilibrium freight flow assignment pattern in an explicit form for the network with one consumer-supplier pair, a single layer of hubs, and affine performance functions. The findings of the paper can give fresh managerial insights; in particular, we show risks arising within available pricing strategies.

Keywords: Nonlinear optimization · Freight flow assignment · Intermodal transportation

1 Introduction

Recent decades have demonstrated the trend towards the use of intermodal freight transportation [12]. Delivery via an intermodal logistics network has appeared to result in lower costs and less congestion than observed with the most prevalent mode of transportation [1]. As a result, intermodal logistics networks have gained importance as a research area due to their positive influence on transportation economics. Indeed, one key strategic planning problem in intermodal freight transportation concerns the design of its logistics network [15]. Decisions made at this level of planning have an impact on the physical infrastructure network and necessitate large capital investments over long time

The work was supported by a grant from the Russian Science Foundation (No. 22-71-10063 Development of intelligent tools for optimization multimodal flow assignment systems in congested networks of heterogeneous products).

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 100–111, 2024.
https://doi.org/10.1007/978-3-031-53552-9_9

horizons. Researchers identify the following strategic level design problems suitable for intermodal transport: hub location problem, network design problem, and regional multimodal planning problem [5]. Handling these problems helps decision-makers verify the impact of infrastructure modifications, the evolution of demand, or government and industry policies [4]. Moreover, planning problems in intermodal freight transport can be related to four types of decision makers, based on the four main activities in intermodal freight transport [3]: drayage operators [9], terminal operators [6], network operators [13], and intermodal hub operators [7]. In this research, we concentrate on the intermodal network level, taking into account hub delays [11] and movement costs [8].

Different approaches have been used to handle intermodal transportation services [2]. In this research, we consider a network topology that combines connected hubs with point-to-point shipping for intermodal transportation. In other words, commodities can be shipped directly from an origin to a destination, or they can be moved to an intermediate hub or terminal [10]. At the hub, the operator consolidates commodities and forwards them to another hub or destination. While one can obtain valuable insights by using a constant transportation cost, there is a need for a more accurate cost function to make the mathematical formulation more applicable in real world cases. In the configuration of this paper, the larger flows at hubs may lead to congestion that influences the final transportation costs. The basic assumption of this study is that in the long run, the activities carried out by shippers will be in equilibrium, i.e., the cost of any shipment cannot be lowered by changing mode, route, or both. Unlike previous research, we assume that commodities may visit as many hubs as needed to reduce costs for transportation. The model of our research exploits the following equilibrium principle: the cost on all used shipping routes via different modes (road- only, rail-only, and intermodal) is equal for each supplier-consumer pair and equal to or less than the cost on any unused routes [14].

The rest of the paper is organized as follows. Section 2 introduces an intermodal transportation network as a directed graph with hub performance functions. Section 3 contains the formulation of a task for an equilibrium freight flow assignment pattern search as a non-linear optimization problem. Section 4 contains the equilibrium freight flow assignment pattern in an explicit form for the network with one consumer-supplier pair, a single layer of hubs, and affine performance functions. Section 5 discusses managerial insights for the hub operator concerning operation pricing strategies. Section 6 is the conclusion.

2 Intermodal Transportation Network

In this paper, we consider the assignment of heterogeneous product flows in the intermodal transportation network, presented by directed graph $G = (V, E)$. A set of nodes V is the union of subsets of consumers V_c, suppliers V_s, and logistic hubs V_h, while

$$V_c = \cup_{w \in W} V_c^w \quad \text{and} \quad V_s = \cup_{w \in W} V_s^w,$$

where W is the set of different types of products. A set of arcs E represents the transport connectivity of the network as a set of available paths for direct shipment from one node to another. For any $w \in W$, a set of nodes sequentially linked by arcs initiating at the node of the supplier from V_s^w and terminating at the node of the consumer from V_c^w is what we call the shipping *route*. The ordered set of all possible shipping routes to consumer $\nu \in V_c^w$, $w \in W$, from all suppliers, included in V_s^w, is denoted as R^ν. Every consumer $\nu \in V_c$ is associated with demand $d_\nu > 0$ (in units of intermodal containers), which may be satisfied by using available shipping routes R^ν, i.e.,

$$\sum_{r \in R^\nu} f_r^\nu = d_\nu \quad \forall \nu \in V_c,$$

where $f_r^\nu \geq 0$ is a variable freight flow through the shipping route $r \in R^\nu$; $f = (\ldots, f_r^\nu, \ldots)^{\mathrm{T}}$.

We use $x_u \geq 0$, $u \in V_s$, to denote the variable load (determined by the volume of production orders) of supplier u:

$$x_u = \sum_{\nu \in V_c} \sum_{r \in R^\nu} f_r^\nu \delta_{u,r}^\nu \quad \forall u \in V_s, \tag{1}$$

where

$$\delta_{u,r}^\nu = \begin{cases} 1, \text{ if } u \in V_s \text{ is initiating node of route } r, \\ 0, \text{ otherwise }, \end{cases} \quad \forall r \in R^\nu, \nu \in V_c.$$

We use $x_v \geq 0$, $v \in V$, to denote the variable load (determined by the volume of handling cargo) of hub v:

$$x_v = \sum_{\nu \in V_c} \sum_{r \in R^\nu} f_r^\nu \delta_{v,r}^\nu \quad \forall v \in V_h, \tag{2}$$

where

$$\delta_{v,r}^\nu = \begin{cases} 1, \text{ if } v \in V_h \text{ belongs to route } r, \\ 0, \text{ otherwise }, \end{cases} \quad \forall r \in R^\nu, \nu \in V_c.$$

By $x_e \geq 0$, $e \in E$, we denote the variable shipment flow through arc e:

$$x_e = \sum_{\nu \in V_c} \sum_{r \in R^\nu} f_r^\nu \delta_{e,r}^\nu \quad \forall e \in E, \tag{3}$$

where

$$\delta_{e,r}^\nu = \begin{cases} 1, \text{ if } e \in E \text{ belongs to route } r, \\ 0, \text{ otherwise }, \end{cases} \quad \forall r \in R^\nu, \nu \in V_c.$$

Moreover, we denote x as the following vector $x = (\ldots, x_u, \ldots, x_v, \ldots, x_e, \ldots)^{\mathrm{T}}$.

We also introduce scalar-valued functions $r_u(x_u)$, $h_v(x_v)$, and $t_e(x_e)$ of class C^1, for $u \in V_s$, $v \in V_h$, and $e \in E$. We suppose that introduced functions are non-negative, and their first derivatives are strictly positive on the set of real non-negative numbers. Function $r_u(x_u)$ reflects a realization price of the

supplier u, $u \in V_s$, which naturally depends on its load by production orders. Function $h_v(x_v)$ describes operation costs in hub v, $v \in V$, which depend on the load of cargo handled. In turn, function $t_e(x_e)$, $e \in E$, describes transportation costs on arc e. In this study, we assume that the *purchase* costs with respect to the shipping route are the sum of the realization price of the supplier, hub operation costs at all hubs belonging to this route, and transportation costs on all arcs belonging to this route. In other words, we define the purchase costs with respect to the shipping route $r \in R^\nu$, $\nu \in V_c$, as the following additive function:

$$p_r^\nu(f) = \sum_{u \in V_s} r_u(x_u)\delta_{u,r}^\nu + \sum_{v \in V_h} h_v(x_v)\delta_{v,r}^\nu + \sum_{e \in E} t_e(x_e)\delta_{e,r}^\nu. \tag{4}$$

3 Freight Flow Assignment

Let us consider the *freight flow assignment* problem in the intermodal transportation network, formulated as follows:

$$\min_x \sum_{u \in V_s} \int_0^{x_u} r_u(\omega)d\omega + \sum_{v \in V_h} \int_0^{x_v} h_v(\omega)d\omega + \sum_{e \in E} \int_0^{x_e} t_e(\omega)d\omega \tag{5}$$

subject to

$$\sum_{r \in R^\nu} f_r^\nu = d_\nu \quad \forall \nu \in V_c, \tag{6}$$

$$f_r^\nu \geq 0 \quad \forall r \in R^\nu, \nu \in V_c, \tag{7}$$

under definitional constraints (1)–(3).

Proposition 1. *The solution $\hat{x}$ to (5)–(7) is unique, and for any $\nu \in V_c$ there exists π^ν such that*

$$p_r^\nu(\hat{f}) \begin{cases} = \pi^\nu, \text{ if } \hat{f}_r^\nu > 0, \\ \geq \pi^\nu, \text{ if } \hat{f}_r^\nu = 0, \end{cases} \quad \forall r \in R^\nu, \nu \in V_c, \tag{8}$$

where $\hat{f}$ satisfies (6)–(7), while $\hat{x}$ and $\hat{f}$ satisfy (1)–(3).

Proof. Since functions $r_u(x_u)$, $h_v(x_v)$, and $t_e(x_e)$ of class C^1, for $u \in V_s$, $v \in V_h$, and $e \in E$, are strictly increasing, their integrals with a variable upper limit are convex functions. Thus, problem (5)–(7) has convex goal function and convex constraints. Consequently, the solution $\hat{x}$ to (5)–(7) is unique. Moreover, there exists at least one pattern $\hat{f}$ such that $\hat{x}$ and $\hat{f}$ satisfy (1)–(3).

Let us study the Lagrangian function of problem (5)–(7):

$$L = \sum_{u \in V_s} \int_0^{x_u} r_u(\omega)d\omega + \sum_{v \in V_h} \int_0^{x_v} h_v(\omega)d\omega + \sum_{e \in E} \int_0^{x_e} t_e(\omega)d\omega + \\ + \sum_{\nu \in V_c} \pi^\nu \left(d_\nu - \sum_{r \in R^\nu} f_r^\nu \right) + \sum_{\nu \in V_c} \sum_{r \in R^\nu} (-f_r^\nu)\xi_r^\nu,$$

where π^ν, $\nu \in V_c$, and $\xi_r^\nu \geq 0$, $r \in R^\nu$, $\nu \in V_c$, are Lagrange multipliers. Since $\hat{f}$ is the route-flow solution to (5)–(7), then $\hat{f}$ has to satisfy Karush-Kuhn-Tucker conditions. First of all, the stationarity should be hold:

$$\left.\frac{\partial L}{\partial f_r^\nu}\right|_{f=\hat{f}} = 0 \quad \forall r \in R^\nu, \nu \in V_c. \tag{9}$$

The first partial derivative of L with respect to f_r^ν, $r \in R^\nu$, $\nu \in V_c$, is

$$\frac{\partial L}{\partial f_r^\nu} = \frac{\partial}{\partial f_r^\nu}\left(\sum_{u \in V_s}\int_0^{x_u} r_u(\omega)d\omega\right) + \frac{\partial}{\partial f_r^\nu}\left(\sum_{v \in V_h}\int_0^{x_v} h_v(\omega)d\omega\right) +$$
$$+ \frac{\partial}{\partial f_r^\nu}\left(\sum_{e \in E}\int_0^{x_e} t_e(\omega)d\omega\right) - \pi^\nu - \xi_r^\nu \quad \forall r \in R^\nu, \nu \in V_c.$$

The first summand of $\partial L/\partial f_r^\nu$ can be re-written as follows:

$$\frac{\partial}{\partial f_r^\nu}\left(\sum_{u \in V_s}\int_0^{x_u} r_u(\omega)d\omega\right) = \sum_{u \in V_s}\frac{\partial}{\partial f_r^\nu}\left(\int_0^{x_u} r_u(\omega)d\omega\right) \quad \forall r \in R^\nu, \nu \in V_c,$$

or

$$\sum_{u \in V_s}\frac{\partial}{\partial f_r^\nu}\left(\int_0^{x_u} r_u(\omega)d\omega\right) = \sum_{u \in V_s}\frac{\partial}{\partial x_u}\left(\int_0^{x_u} r_u(\omega)d\omega\right)\frac{\partial x_u}{\partial f_r^\nu} \quad \forall r \in R^\nu, \nu \in V_c.$$

However, according to (1),

$$\frac{\partial x_u}{\partial f_r^\nu} = \delta_{u,r}^\nu \quad \forall u \in V_s, r \in R^\nu, \nu \in V_c,$$

hence,

$$\frac{\partial}{\partial f_r^\nu}\left(\sum_{u \in V_s}\int_0^{x_u} r_u(\omega)d\omega\right) = \sum_{u \in V_s} r_u(x_u)\delta_{u,r}^\nu \quad \forall r \in R^\nu, \nu \in V_c.$$

The second summand of $\partial L/\partial f_r^\nu$ can be re-written as follows:

$$\frac{\partial}{\partial f_r^\nu}\left(\sum_{v \in V_h}\int_0^{x_v} h_v(\omega)d\omega\right) = \sum_{v \in V_h}\frac{\partial}{\partial f_r^\nu}\left(\int_0^{x_v} h_v(\omega)d\omega\right) \quad \forall r \in R^\nu, \nu \in V_c,$$

or

$$\sum_{v \in V_h}\frac{\partial}{\partial f_r^\nu}\left(\int_0^{x_v} h_v(\omega)d\omega\right) = \sum_{v \in V_h}\frac{\partial}{\partial x_v}\left(\int_0^{x_v} h_v(\omega)d\omega\right)\frac{\partial x_v}{\partial f_r^\nu} \quad \forall r \in R^\nu, \nu \in V_c.$$

However, according to (2),

$$\frac{\partial x_v}{\partial f_r^\nu} = \delta_{v,r}^\nu \quad \forall v \in V_h, r \in R^\nu, \nu \in V_c,$$

hence,

$$\frac{\partial}{\partial f_r^\nu}\left(\sum_{v\in V_h}\int_0^{x_v} h_v(\omega)d\omega\right)=\sum_{v\in V_h} h_v(x_v)\delta_{v,r}^\nu \quad \forall r\in R^\nu,\nu\in V_c.$$

The third summand of $\partial L/\partial f_r^\nu$ can be re-written as follows:

$$\frac{\partial}{\partial f_r^\nu}\left(\sum_{e\in E}\int_0^{x_e} t_e(\omega)d\omega\right)=\sum_{e\in E}\frac{\partial}{\partial f_r^\nu}\left(\int_0^{x_e} t_e(\omega)d\omega\right) \quad \forall r\in R^\nu,\nu\in V_c,$$

or

$$\sum_{e\in E}\frac{\partial}{\partial f_r^\nu}\left(\int_0^{x_e} t_e(\omega)d\omega\right)=\sum_{e\in E}\frac{\partial}{\partial x_e}\left(\int_0^{x_e} t_e(\omega)d\omega\right)\frac{\partial x_e}{\partial f_r^\nu} \quad \forall r\in R^\nu,\nu\in V_c.$$

However, according to (3),

$$\frac{\partial x_e}{\partial f_r^\nu}=\delta_{e,r}^\nu \quad \forall e\in E, r\in R^\nu,\nu\in V_c,$$

hence,

$$\frac{\partial}{\partial f_r^\nu}\left(\sum_{e\in E}\int_0^{x_e} t_e(\omega)d\omega\right)=\sum_{e\in E} t_e(x_e)\delta_{e,r}^\nu \quad \forall r\in R^\nu,\nu\in V_c.$$

Therefore, for all $r\in R^\nu$, $\nu\in V_c$,

$$\frac{\partial L}{\partial f_r^\nu}=\sum_{u\in V_s} r_u(x_u)\delta_{u,r}^\nu+\sum_{v\in V_h} h_v(x_v)\delta_{v,r}^\nu+\sum_{e\in E} t_e(x_e)\delta_{e,r}^\nu-\pi^\nu-\xi_r^\nu,$$

or, according to (4),

$$\frac{\partial L}{\partial f_r^\nu}=p_r^\nu(f)-\pi^\nu-\xi_r^\nu \quad \forall r\in R^\nu,\nu\in V_c,$$

that, due to (9), leads to

$$p_r^\nu(\hat{f})=\pi^\nu+\xi_r^\nu \quad \forall r\in R^\nu,\nu\in V_c.$$

Moreover, according to primal feasibility, $\hat{f}$ satisfies (6)–(7), while, according to complementary slackness, $(-f_r^\nu)\xi_r^\nu=0$ for all $r\in R^\nu$, $\nu\in V_c$, and, according to dual feasibility, $\xi_r^\nu\geq 0$ for all $r\in R^\nu$, $\nu\in V_c$. Consequently,

$$p_r^\nu(\hat{f})\begin{cases}=\pi^\nu, \text{ if } \hat{f}_r^\nu>0\\ \geq\pi^\nu, \text{ if } \hat{f}_r^\nu=0\end{cases} \quad \forall r\in R^\nu,\nu\in V_c,$$

where $\hat{f}$ satisfies (6)–(7), while $\hat{x}$ and $\hat{f}$ satisfy (1)–(3). □

Therefore, the solution to (5)–(7) gives one an estimation of the equilibrium freight flow assignment pattern in the intermodal transportation network, presented by graph G.

4 A Single Layer of Hubs in the Case of Affine Functions

In this section, we consider an intermodal transportation network with a single layer of hubs. The example of such a network is given in Fig. 1. This network has one supplier (node 1), one consumer (node 6), and a single layer of hubs (nodes 2–5).

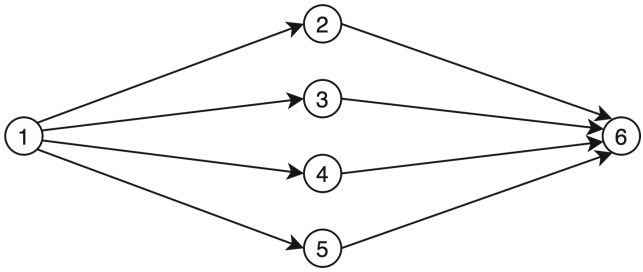


Fig. 1. The intermodal transportation network with a single layer of hubs

In the general case of the considered network, the demand of a single consumer d seeks to be assigned among n available routes: $d = \sum_{i=1}^{n} f_i$; while the realization price of an supplier is $p_1(d) = p^0$. The hub performance function determines the hub operation pricing and has the following form: $h_i(f_i) = h_i^0 + \alpha_i f_i$, for all $i = \overline{1,n}$. Transportation costs on the route i, $i = \overline{1,n}$, can be computed due to the affine function $t_i(f_i) = t_i^0 + f_i/c_i$, where c_i is the capacity of the route, for all $i = \overline{1,n}$.

Without loss of generality, we assume that

$$h_1^0 + t_1^0 \leq \ldots \leq h_n^0 + t_n^0 \tag{10}$$

Proposition 2. *In the case of an intermodal transportation network with one consumer-supplier pair, a single layer of hubs, and affine functions, the solution* $\hat{f}$ *to* (5)–(7) *has the following form:*

$$\hat{f}_i = \frac{1}{\alpha_i + \frac{1}{c_i}} \frac{d + \sum_{j=1}^{k} \frac{r_0 + h_j^0 + t_j^0}{\alpha_j + \frac{1}{c_j}}}{\sum_{j=1}^{k} \frac{1}{\alpha_j + \frac{1}{c_j}}} - \frac{r_0 + h_i^0 + t_i^0}{\alpha_i + \frac{1}{c_i}}, \quad \forall i = \overline{1,k},$$

where k *satisfies inequalities:*

$$r^0 + h_k^0 + t_k^0 < \frac{d + \sum_{j=1}^{k} \frac{r_0 + h_j^0 + t_j^0}{\alpha_j + \frac{1}{c_j}}}{\sum_{j=1}^{k} \frac{1}{\alpha_j + \frac{1}{c_j}}} \leq r^0 + h_{k+1}^0 + t_{k+1}^0,$$

while $\hat{f}_i = 0$ *for* $i = \overline{k+1,n}$.

Proof. According to Proposition 1, the solution to (5)–(7) is unique, and for the network with a single layer of hubs and one consumer-supplier pair, there exists π such that

$$r_0 + h_i(\hat{f}_i) + t_i(\hat{f}_i) \begin{cases} = \pi, \text{ if } \hat{f}_i > 0 \\ \geq \pi, \text{ if } \hat{f}_i = 0 \end{cases} \quad \forall i = \overline{1,n}.$$

Since functions $h_i(f_i)$ and $t_i(f_i)$, $i = \overline{1,n}$, are affine functions, then

$$r_0 + h_i^0 + \alpha_i \hat{f}_i + t_i^0 + \frac{\hat{f}_i}{c_i} \begin{cases} = \pi, \text{ if } \hat{f}_i > 0 \\ \geq \pi, \text{ if } \hat{f}_i = 0 \end{cases} \quad \forall i = \overline{1,n},$$

or

$$r_0 + h_i^0 + t_i^0 + \left(\alpha_i + \frac{1}{c_i}\right) \hat{f}_i \begin{cases} = \pi, \text{ if } \hat{f}_i > 0 \\ \geq \pi, \text{ if } \hat{f}_i = 0 \end{cases} \quad \forall i = \overline{1,n}. \tag{11}$$

Due to (10), there exists k such that

$$r_0 + h_1^0 + t_1^0 \leq \ldots r_0 + h_k^0 + t_k^0 < \pi \leq r_0 + h_{k+1}^0 + t_{k+1}^0 \ldots \leq r_0 + h_n^0 + t_n^0,$$

that, according to (11), leads to

$$\hat{f}_i = \frac{\pi - \left(r_0 + h_i^0 + t_i^0\right)}{\alpha_i + \frac{1}{c_i}}, \quad \forall i = \overline{1,k},$$

while $\hat{f}_i = 0$ for $i = \overline{k+1,n}$. Moreover,

$$\sum_{i=1}^{n} \hat{f}_i = \sum_{i=1}^{k} \hat{f}_i = \sum_{i=1}^{k} \frac{\pi - \left(r_0 + h_i^0 + t_i^0\right)}{\alpha_i + \frac{1}{c_i}} = d,$$

and, hence,

$$\pi = \frac{d + \sum_{i=1}^{k} \frac{r_0 + h_i^0 + t_i^0}{\alpha_i + \frac{1}{c_i}}}{\sum_{i=1}^{k} \frac{1}{\alpha_i + \frac{1}{c_i}}}.$$

Eventually, we obtain

$$f_i = \frac{1}{\alpha_i + \frac{1}{c_i}} \frac{d + \sum_{j=1}^{k} \frac{r_0 + h_j^0 + t_j^0}{\alpha_j + \frac{1}{c_j}}}{\sum_{j=1}^{k} \frac{1}{\alpha_j + \frac{1}{c_j}}} - \frac{r_0 + h_i^0 + t_i^0}{\alpha_i + \frac{1}{c_i}}, \quad \forall i = \overline{1,k}.$$

□

5 Strategies for Intermodal Hub Operation Pricing

In the case of an intermodal transportation network with one consumer-supplier pair, a single layer of hubs, and affine functions, according to Proposition 2, the solution $\hat{f}$ to (5)–(7) has the following form:

$$\hat{f}_i = \frac{\pi - r_0 - h_i^0 - t_i^0}{\alpha_i + \frac{1}{c_i}}, \quad \forall i = \overline{1,k}, \tag{12}$$

where

$$\pi = \frac{d + \sum_{j=1}^{k} \frac{r_0 + h_j^0 + t_j^0}{\alpha_j + \frac{1}{c_j}}}{\sum_{j=1}^{k} \frac{1}{\alpha_j + \frac{1}{c_j}}}, \tag{13}$$

while $\hat{f}_i = 0$ for $i = \overline{k+1, n}$. Moreover, according to Proposition 1, $p_i(\hat{f}_i) = \pi$ if $\hat{f}_i > 0$, $i = \overline{1, n}$, i.e., the average purchase costs per a container is equal to π as well. Therefore, by varying α_i for different transportation hubs, one can identify how the hub operation pricing influences the average purchase cost and the hub load.

In this section, we consider the intermodal transportation network with one consumer-supplier pair, a single layer of four hubs, and affine functions: $h_i(f_i) = h_i^0 + \alpha_i f_i$ and $t_i(f_i) = t_i^0 + f_i/c_i$, for all $i = \overline{1,4}$. We assume that $d = 20$ and $c_i = 5$, $i = \overline{1,4}$, while $r^0 = 10$, $h_1^0 = 9$, $h_2^0 = 10$, $h_3^0 = 11$, $h_4^0 = 12$, and $t_1^0 = 5$, $t_2^0 = 6$, $t_3^0 = 7$, $t_4^0 = 8$. Firstly, we vary α_i in the segment $[0.1, 100]$ separately for every i, $i = \overline{1,4}$, to identify how the hub operation pricing influences the average purchase cost (13). The dependence of the average purchase cost (13) on the hub pricing strategy is given in Fig. 2.

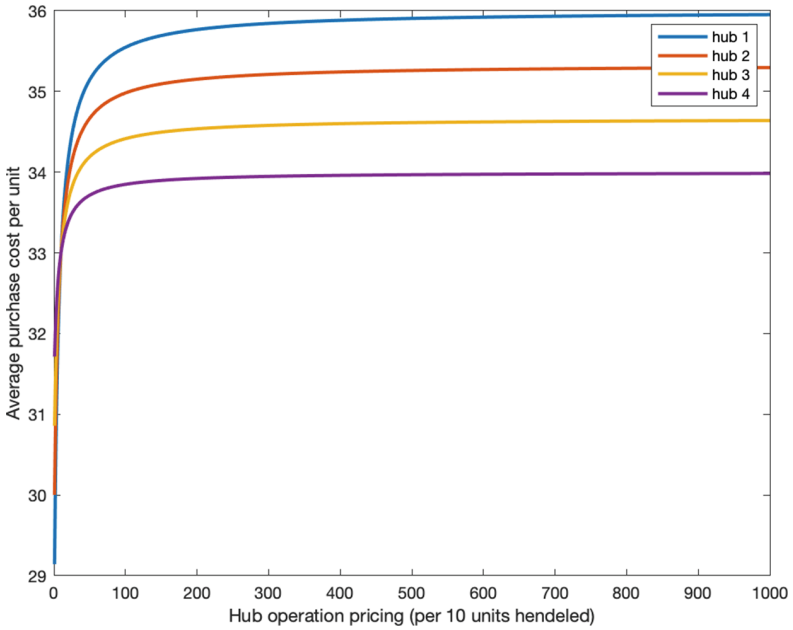


Fig. 2. The dependence of the average purchase cost on the hub pricing strategy

On the one hand, the value of the average purchase cost is most sensitive to the pricing strategy of the hub with the lowest value of $h_i^0 + t_i^0$, $i = \overline{1,4}$.

Moreover, the higher the value of $h_i^0+t_i^0$, $i=\overline{1,4}$, the less the value of the average purchase cost caused by the pricing strategy of such a hub. In other words, the most influence on the value of the average purchase cost is occurred by the route with the cheapest price of non-congested (flow-free) shipping. Indeed, according to (13), when one changes α_i in the hub with the lowest value of $h_i^0+t_i^0$, he or she influences the lowest-value summand of π (compared to others). On the other hand, the most growth of the average purchase cost under little changes of α_i occurs when $\alpha_i \in [0.1, 10]$, $i=\overline{1,4}$. Therefore, the pricing strategies with respect to the average purchase cost have their limits, i.e., there are segments of α_i for which big changes in α_i lead to small changes in the average purchase cost.

Secondly, we vary α_i in the segment [0.1, 100] separately for every i, $i=\overline{1,4}$, to identify how the hub operation pricing influences the hub load (12). The dependence of the hub load (12) on the hub pricing strategy is given in Fig. 3.

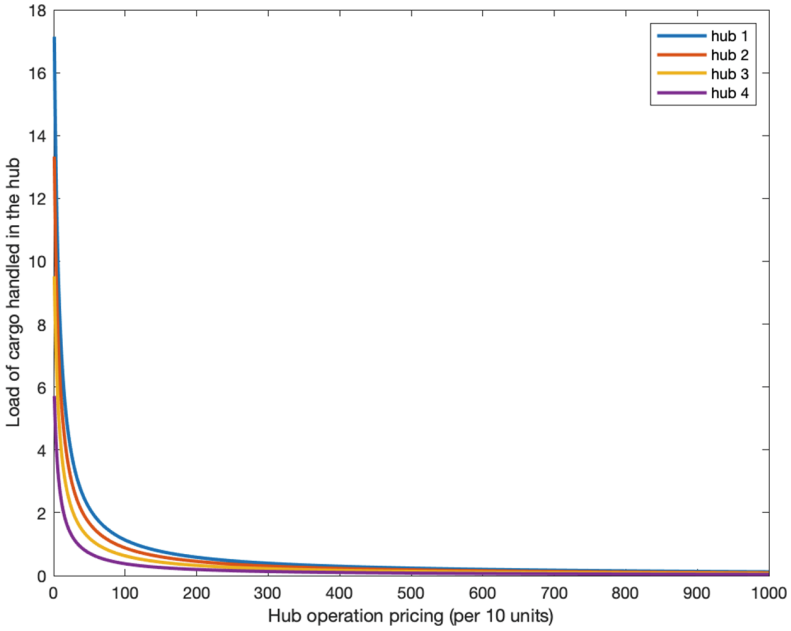


Fig. 3. The dependence of the hub load on the pricing strategy

On the other hand, the value of the hub load is most sensitive to the pricing strategy of the hub with the highest value of $h_i^0+t_i^0$, $i=\overline{1,4}$. In other words, the most influence on the value of the hub load is occurred by the route with the most expensive price of non-congested (flow-free) shipping. On the other hand, the most decrease of the hub load under little changes of α_i occurs when $\alpha_i \in [0.1, 10]$, $i=\overline{1,4}$. Therefore, the pricing strategies with respect to the

average purchase cost have their limits, i.e., there are segments of α_i for which big changes in α_i lead to small changes in the hub load.

Therefore, if the hub pricing manager follows a decreasing (increasing) strategy in order to avoid the underload (overload), he or she can face several risks raised as side effects of this strategy. The first risk is the weak return. In other words, α_i can be located in such a segment that big changes in α_i lead to little changes in the hub load. The second risk is a drastic increase in the average purchase cost. Indeed, α_i can be located in such a segment that little changes in α_i lead to big changes in the average purchase cost (Table 1).

Table 1. Strategies for the hub operation pricing.

Hub status	Strategy	Risk
Underload	Decrease α_i to reduce the contribution of the hub operation costs in purchase costs	Weak return
Overload	Increase α_i to intensify the contribution of the hub operation costs in purchase costs	Drastic increase in the average purchase cost

6 Conclusion

This work was focused on hub operation pricing under an equilibrium freight flow assignment model. We formulated this model as a non-linear optimization problem and showed that its solution corresponds to the equilibrium assignment pattern in an intermodal transportation network. We analyzed the sensitivity of the average purchase cost and hub load to different strategies of hub operation pricing. To this end, we obtained the equilibrium freight flow assignment pattern in an explicit form for the network with one consumer-supplier pair, a single layer of hubs, and affine performance functions. The findings of the paper can give fresh managerial insights; in particular, we showed risks arising within available pricing strategies.

References

1. Agamez-Ariasa, A.M., Moyano-Fuentesb, J.: Intermodal transport in freight distribution: a literature review. Transp. Rev. **37**(6), 782–807 (2017)
2. Alumur, S., Kara, B.Y.: Network hub location problems: the state of the art. Eur. J. Oper. Res. **190**, 1–21 (2008)
3. Caris, A., Macharis, C., Janssens, G.K.: Decision support in intermodal transport: a new research agenda. Comput. Ind. **64**, 105–112 (2013)
4. Caris, A., Macharis, C., Janssens, G.K.: Planning problems in intermodal freight transport: accomplishments and prospects. Transp. Plan. Technol. **31**(3), 277–302 (2008)

5. Crainic, T.G., Laporte, G.: Planning models for freight transportation. Eur. J. Oper. Res. **97**, 409–438 (1997)
6. Corry, P., Kozan, E.: An assignment model for dynamic load planning of intermodal trains. Comput. Oper. Res. **33**, 1–17 (2006)
7. Erera, A.L., Morales, J.C., Savelsbergh, M.: Global intermodal tank container management for the chemical industry. Transp. Res. Part E **41**, 551–566 (2005)
8. Ghane-Ezabadi, M., Vergara, H.A.: Integrated intermodal network design with nonlinear inter-hub movement costs. In: 15th IMHRC Proceedings. Savannah, Georgia, USA (2018)
9. Imai, A., Nishimura, E., Current, J.: A Lagrangian relaxation-based heuristic for the vehicle routing with full container load. Eur. J. Oper. Res. **176**(1), 87–105 (2007)
10. Ishfaq, R., Sox, C.R.: Hub location-allocation in intermodal logistic networks. Eur. J. Oper. Res. **210**, 213–230 (2011)
11. Ishfaq, R., Sox, C.R.: Design of intermodal logistics networks with hub delays. Eur. J. Oper. Res. **220**, 629–641 (2012)
12. Macharis, C., Bontekoning, Y.M.: Opportunities for OR in intermodal freight transport research: a review. Eur. J. Oper. Res. **153**, 400–416 (2004)
13. Racunica, I., Wynter, L.: Optimal location of intermodal freight hubs. Transp. Res. Part B **39**, 453–477 (2005)
14. Wardrop, J.G.: Some theoretical aspects of road traffic research. Proc. Inst. Civ. Eng. **1**(3), 325–362 (1952)
15. Woxenius, J.: Generic framework for transport network designs: applications and treatment in intermodal freight transport literature. Transp. Rev. **27**, 733–749 (2007)

A New Approach to Eliminating the "Flip" Effect of the Approximating Function Under Conditions of a Priori Uncertainty

V. I. Marchuk(✉), A. A. Samohleb, M. A. Laouar, and H. T. A. Al-Ali

Don State Technical University, 1, Gagarin Square, Rostov-On-Don 344000, Russia
marchuk@sssu.ru

Abstract. The paper considers a new approach to compensate for errors that arise when processing measurement results under conditions of a priori uncertainty due to the appearance of the "flip" effect when approximating the measured function over a certain observation interval. The proposed approach is to increase the approximation interval at which a "flip" is detected to a value at which this "flip" effect is not observed. However, the obtained approximation values are used only on the interval of the initial value, i.e. the interval where the "flip" effect was observed. The approach used for processing the measurement results makes it possible to significantly reduce the error in estimating the function of the measured process under conditions of a priori uncertainty. #COMESYSO1120.

Keywords: Approximation · Error Rate · Minimizing Errors · Mirroring Function

1 Introduction

A large number of experimental data obtained during scientific and industrial experiments requires their processing under conditions of a limited amount of a priori information about the measured process and statistical characteristics of the additive noise component. The main difficulty in separating the initial implementation of measurement results into useful and random components is the lack of a priori information about the statistics of these processes and the requirement for full automation of the procedures for selecting components without the participation of an experimenter. Using in practice the methods that are considered in the works of David G., Perevertkin S.M., Bendat J., Andersen T., Wiener N., Kalman R.E., Brandt Z., Levin B.R., Tsvetkov E.I., Tikhonov V.I., Krinetsky E.I., Fomin A.F., Ayvazyan S.A., Likharev V.A. and a number of others, it is possible only if there is the necessary amount of a priori information about both useful and random components. Otherwise, their effectiveness decreases, and it becomes impractical to talk about the reliability of the analysis. This explains the fact that with such a huge amount of work on this topic, computer data processing is carried out either with additional visual analysis, or with the simplest methods, such as the moving average method and its modifications. This is especially evident when processing measurement

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 112–118, 2024.
https://doi.org/10.1007/978-3-031-53552-9_10

results, which are represented by a single implementation of the measured process. In this regard, the solution of the task in conditions of a limited amount of a priori information is extremely relevant and has high scientific and practical significance.

Currently, as the analysis of literature sources [1–17] shows, the method of multiplication of estimates (RAZOC), which is described quite well in [10–13, 18–21], has the highest efficiency in conditions of a limited amount of a priori information about the function of the measured process and statistical characteristics of the additive noise component.

The principle of the method involves splitting the original implementation into some intervals using a random number sensor, followed by approximation by a polynomial of the second degree. Thus, the first estimate of the measured process is obtained. The principle of the method involves splitting the original implementation into some intervals using a random number sensor, followed by approximation by a polynomial of the second degree. Thus, the first estimate of the measured process is obtained. This procedure is repeated P times, resulting in P estimates of the measured process. Then, at each moment of time, the value of all estimates is averaged and the resulting estimate of the useful signal is obtained. A more detailed description of the principle of operation of the RAZOC is given in [3, 4, 10, 18, 22].

However, it is noted in [22–24] that when studying the RAZOC method, the effect of a "flip" of the approximating function was detected at some approximation intervals, i.e. if at this interval the approximating polynomial should be described by the expression $y(t) = a_0 + b_0t + c_0t^2$, and it is described by the expression $y(t) = a_0 + b_0t\text{-}c_0t^2$, that is, the direction of the branches of the polynomial changes, then such a change is called the "flip" effect of the approximating function. The presence of this effect significantly increases the error in the allocation of the useful signal function, however, it can only be determined visually at the present level of research, the detection of the "flip" effect in automatic mode will be proposed by the authors in the coming papers. In [23–25], some solutions are proposed to eliminate this effect by mirroring the approximating function in the presence of the "flip" effect. However, the research results given in [23–25] show that reducing the error by using a mirror image of the approximating function on the interval where the "flip" effect is detected allows reducing its effect on the error of the useful signal selection. The analysis of the research results of the RAZOC method with the detection of the "flip" effect of the approximating function and its compensation due to the mirror flip of the approximating function are given in [24, 25].

2 Formulation of the Problem

In this paper, a new method is proposed to eliminate the "flip" of the approximating function on a certain approximation interval without using a mirror image on this interval. In this regard, we will consider a special case of eliminating the effect of the "flip" of the approximating function on one estimate of the useful signal, which is a special case of using the RAZOC method.

Suppose that the initial implementation of the measurement results can be represented as a sum

$$Y(t) = S(t) + \eta(t) \tag{1}$$

where S(t) is the useful signal to be allocated; η(t) – additive noise.

As a result of the approximation of the original function Y(t), we obtain an estimate of the useful signal $\overline{S(t)}$ with a certain standard error σ_1. It was shown in [23] that the probability of the appearance of the "flip" effect strongly depends on the size of the approximation interval and the dispersion of the additive noise component. Suppose that the function of the useful signal is described by a polynomial of the second degree, and the noise dispersion is increased until the "flip" effect appears, (see Fig. 1).

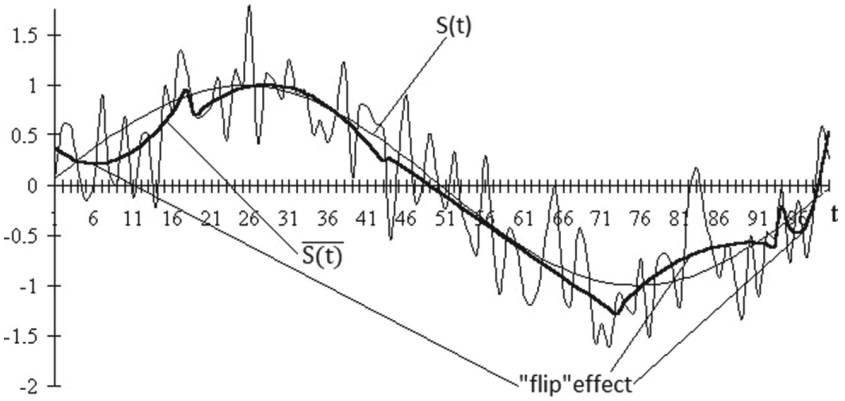


Fig. 1. The result of the appearance of the "flip" effect when approximating the measurement results.

To ensure that the research results do not depend on the structure of the additive noise used, we will use the average of the RMS error of the approximation results obtained for each of the 1000 implementations without the presence of a "flip" and which is equal to $m_{\sigma 1} = 0.002476$ and denote it as σ_2. By changing the noise implementation at its constant variance, we will find implementations in which there is a "flip" of the approximating function, the standard error in this case will be equal to σ_3, and the average value of the standard error σ_3 of the approximation results for 1000 implementations with a "flip" is equal to $m_{\sigma 3} = 0.352567$, which is significantly higher than $m_{\sigma 1}$ (more than 140 times). Using to eliminate the "flip" effect by its mirror image according to the method presented in [25], the value of the root-mean-square error is σ_4, and the average value of the root-mean-square error of the approximation results for 1000 implementations when eliminating the "flip" effect by its mirror image is $m_{\sigma 4} = 0.255049$, which is 27% less than $m_{\sigma 3}$ when the presence of a " flip ".

3 Discussion of Research Results

In this paper, a new method is proposed to reduce the root-mean-square error of the approximation results in the presence of the "flip" effect, which consists of the following:

– A sample of length N is taken, within which a certain interval N_1 is selected and by changing the noise structure at a given variable, we achieve the effect of a "flip" on the interval N_1;

– The approximation interval N_1 is determined, at which the "flip" effect is observed;
– For a given sample N, the interval N_1 increases to the value N_2, until the "flip" effect ceases to be observed [25];
– An approximation is carried out on the interval N_2, however, the result of the approximation is replaced with the original sample only on the interval N_1, i.e. on the interval where the "flip" effect was observed.

In the work, studies were carried out on the probability of eliminating the " flip " effect with an increase in the approximation interval N_2 according to the proposed method. The results of the studies are shown in Fig. 2. The studies were carried out at intervals $N_1 = 5$ for 1000 realizations of plots with an increase in the interval N_2 by 2, 4, 6, 8 and 10 in relation to the value of the interval N_1 values, followed by the calculation of the probability of the absence of a " flip " in the interval N_2.

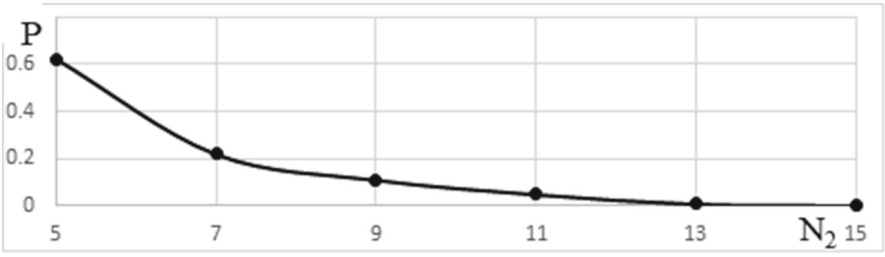


Fig. 2. The probability of a "flip" with an increase in the interval N_1 with a noise dispersion of 0.1.

An analysis of the results presented in Fig. 2 shows that increasing the sample even by 2 values, i.e. $N_2 = N_1 + 2$, reduces the probability of a " flip " effect by 2 times. With an increase in the sample of $N_2 = N_1 + 8$ values, the probability of a " flip " effect decreases to almost zero. Thus, the assumption that an increase in the sample length during approximation can reduce or completely eliminate the "flip" effect is experimentally confirmed.

Let's consider the change in the root-mean-square error of allocating a useful signal without the presence of the "flip" effect - $m\sigma 1$, in the presence of the "flip" effect – $m_{\sigma 3}$, as well as when eliminating the "flip" effect according to the proposed method - $m_{\sigma 5}$, where σ_5 is the root-mean-square error for one implementation when eliminating the "flip" effect according to the proposed method. The studies were carried out with a different number of observed N intervals with the presence of the " flip " effect. The results of the conducted studies are shown in Fig. 3

As can be seen from the presented research results in Fig. 3, with an increase in the number of "overturns" on one sample with an interval of $N_1 = 5$, the error value $m\sigma 3$ increases almost linearly (curve 2), while the increase in the standard error $m\sigma 3$ varies from 5 to 14 times depending on the number of "overturns" in the sample under study. At the same time, it should be noted that curve 1 shows the approximation error of the sample $m\sigma 1$ in the absence of the "flip" effect in the sample under study.

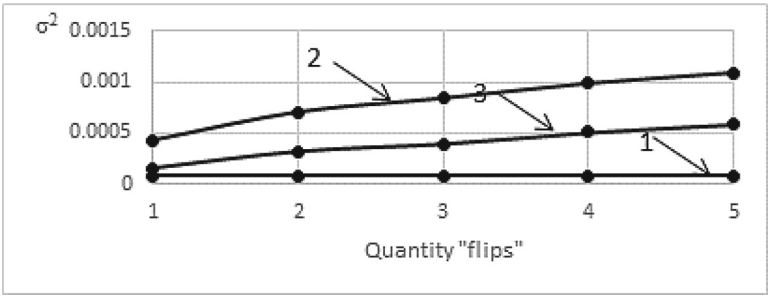


Fig. 3. Dependence of the change in the standard error in the presence of the "flip" and its elimination.

The use of the proposed methodology for processing measurement results under conditions of a priori uncertainty (curve 3) makes it possible to reduce the standard error $m_{\sigma 5}$ by an average of 2 times compared with the presence of "reversals" $m_{\sigma 3}$ (curve 2). It should be noted that the use of mirroring to eliminate the "flip" effect reduces the processing error, as mentioned above, with an average of 27%.

Thus, the results of the conducted studies have shown that the use of the proposed technique is an order of magnitude more effective than the mirroring method with the effect of the "flip" of the approximating function.

4 Conclusions

The conducted studies on the evaluation of the proposed method of processing measurement results under conditions of a priori uncertainty have shown that:

Using the mirror reflection method to eliminate the "flip" effect gives a decrease in the RMS error value of about 27%.

Increasing the sample even by 2 values reduces the probability of a "flip" effect by 2 times. When the sample is increased to 8 values, the probability of a "flip" effect decreases to almost zero.

With an increase in the number of "flips" on one sample with an interval of $N_1 = 5$, the error value increases almost linearly.

The proposed method of processing measurement results under conditions of a priori uncertainty allows reducing the standard error by an average of 2 times compared with the error in the presence of "flips".

References

1. Perevertkin, S.M., Kantor, A.V., Borodin, N.F., Shcherbakova, T.S.: On-board telemetry equipment of space aircraft. Machinostroenie, Moscow (1977)
2. Borovikov, V.P.: The art of data analysis on a computer. For professionals. Statistika St. Petersburg (2001)
3. Marchuk, V.I.: Increasing the reliability of the primary processing of the measurement results. Measur. Equip. **12**, 3–5 (2003)

4. Marchuk, V.I.: Primary processing of measurement results with a limited amount of a priori information. TRSTU, Taganrog (2003)
5. Golyandina, N.E.: The «Caterpillar» method-SSA for analysis of time series. St. Petersburg (2004)
6. Marchuk, V.I., Rumyantsev, K.E.: A new way to increase the reliability of measurement results during rocket and space research. Aerospace Instrum. 2 (2004)
7. Marchuk, V.I., Rumyantsev, K.E., Sherstobitov, A.I.: Filtration of low-frequency processes with a limited volume of measurement results. Radio Eng. **9**, 3–7 (2006)
8. Mishulina, O.A.: Statistical analysis and processing of time series. MEPhI (2008)
9. Sadovnikova, N.A., Shmoylova, R.A.: Time series analysis and forecasting. "Futuris" (2009)
10. Marchuk, V.I.: Estimation of the error of approximation of the useful component when dividing the implementation of measurement results into intervals. Telecommun. **8**, 12–16 (2010)
11. Marchuk, V.I., Voronin, V.V., Sherstobitov, A.I.: Estimation of the error of useful signal extraction during processing under conditions of a limited amount of a priori information. Radio Eng. **9**, 75–82 (2011)
12. Marchuk, V., Makov, S., Timofeev, D., Pismenskova, M., Fisunov, A.: A method of signal estimation error reduction in a priori indeterminacy. In: The Collection: 2015 23rd Telecommunications Forum, pp. 400–403. TELFOR 2015, Serbia (2015)
13. Marchuk, V.: Reducing of noise structure influence on an accuracy of a desired signal extraction. Serbian J. Electrical Eng. **15**, 365–370 (2018)
14. Nielsen, A.: Practical Time Series Analysis. O'Reilly Media, Inc. (2019)
15. Sklyar, A.Y.: Analysis and elimination of the noise component in time series with variable pitch, Cybern. Programm. **1**, 51–59 (2019)
16. Kildishev, G.S., Frenkel, A.A.: Time series analysis and forecasting. URSS, Moscow (2021)
17. Slyusareva, V.A., Budantsev, A.V.: Research and forecasting of time series. Actual Res. **21**(100), 33–37 (2022)
18. Marchuk, V.I., Sahakyan, G.R., Ulanov, A.P.: A method for isolating a trend by multiplying estimates of its single initial realization (RAZOC) and a device for its implementation. Pat. 2207622 Russian Federation, MPK7 G 06 F 17/18. the applicant and patent holder of the South-Russian State University of Economics and Service. 2000127308/09; application 30.10.2000; publ. 27.06.03, Bul. 18. p.14 (2000)
19. Marchuk, V.I., Voronin, V.V., Sherstobitov, A.I.: Estimation of the error of useful signal extraction during processing under conditions of a limited amount of a priori information. Radio Eng., 75– 82 (2011)
20. Marchuk, V.I., Voronin, V.V., Sherstobitov, A.I., Semenishchev, E.A.: Methods of digital signal processing for solving applied problems. Monograph Radio Eng., 128 (2012)
21. Marchuk, V.I.: Estimation of the error of approximation of the useful component when dividing the implementation of measurement results into intervals. Modern Inf. Technol. **19**, 153–159 (2014)
22. Marchuk, V.I., Schreifel, I.S.: Methods of extracting a useful component with a priori uncertainty and a limited volume of measurement results. Monograph. Publishing House of Yurgues, Shakhty (2008)
23. Marchuk, V., Chernyshov, D., Sadrtdinov, I., Minaev, A.: Research of the probability of the "flip" of approximating function during the processing of measurement results. In: E3S Web of Conferences, p. 104. EDP Sciences, France (2019)

24. Marchuk, V., Shrayfel, I., Malcev, I.: Solving the problem of mirroring the signal function with respect to a straight line. In: XV International Scientific-Technical Conference "Dynamics of Technical Systems" (DTS-2019), pp. 11–13. AIP Conference Proceedings: Fundamental Methods of System Analysis, Modeling and Optimization of Dynamic Systems, Rostov-on-Don (2019)
25. Marchuk, V., Hripkov, G., Nikishin, I., Shrivel, I.: Methods for minimizing the error in selecting a useful signal in the presence of the «flip» effect of the approximating function. IOP Conf. Ser.: Mater. Sci. Eng. 1029 (2021)

Modeling the Alienability of an Electronic Document

Alexander V. Solovyev(✉)

Federal Research Center "Computer Science and Control" of the Russian Academy of Scienes, 44/2, Vavilova Street, Moscow 119333, Russia
soloviev@isa.ru

Abstract. The article considers aspects of mathematical modeling of the alienability of an electronic document from the software and hardware environment for its storage. Alienability refers to the possibility of extracting an electronic document from the current hardware and software storage environment and placing it in a new one, where it must also be stored, independent of the storage environment and interpreted. A distinctive feature of the electronic document storage environment is mobility and variability, as well as susceptibility to parametric disturbances. It follows from this that during long-term storage, an electronic document, as a control object, should be maximally stabilized in terms of its main characteristic - safety. To do this, in particular, it should be as independent as possible from the storage environment. As a measure of independence, a mathematical model for assessing the risk of inalienability of a document from the storage environment is proposed. The practical use of the proposed modeling aspects in a number of electronic archive projects, as well as the performed preliminary calculations, give reason to believe that the presented model is adequate for the conditions of the long-term storage problem. #COMESYSO1120.

Keywords: Long-Term Storage · Electronic Document · Document Safety · Alienability · Risk Assessment

1 Introduction

As shown in [1], in the case of long-term storage of electronic documents (hereinafter referred to as ED, EDs), the problem of determining the composition of the information necessary and sufficient for the full interpretation of such a document in the future becomes extremely important.

According to [2, 3], a document is defined as structured information, which is a set of interrelated semantic blocks. Semantic blocks - parts of the document, selected by semantic content. Based on the general definition of a document, an electronic document can be defined as a document whose semantic blocks and the relationships between them are presented in electronic digital form.

In [4], semantic blocks for a long-term storage document are defined and a mathematical model of long-term storage is proposed.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 119–127, 2024.
https://doi.org/10.1007/978-3-031-53552-9_11

However, an important unsolved problem remains to determine how realistically stored ED, even if created according to the correct long-term storage model, can be saved with a sharp change in the hardware and software storage environment. By a sharp change we mean significant changes that occur over a short period of time. Such changes can be:

– failure of technical storage facilities (first of all, media on which electronic documents are stored);
– a significant update of software tools (operating systems, interpretation programs (reading, decoding and visualization) of electronic documents);
– change of cryptographic means and standards of information protection (if such are used for long-term storage).

All these changes lead to the fact that EDs must be extracted from the old hardware and software storage environment and placed in a new one, where they must also be stored and interpreted.

Then by alienation we will understand to the possibility of extracting an electronic document from the current hardware and software storage environment and placing it in a new one, where it must also be stored, independent of the storage environment and interpreted.

As shown in [5], the ED storage environment is always mobile and changeable, subject to parametric disturbances. Therefore, during long-term storage, the ED should be, as a control object, maximally stabilized in terms of the main characteristic - safety. To do this, in particular, it should be as independent as possible from the storage environment.

As a measure of independence, this article proposes a mathematical model for assessing the risk of ED inalienability from the software and hardware storage environment.

Let us consider in more detail the aspects of modeling the alienability of ED.

2 Mathematical Formulation of the Problem of Assessing the Alienability of an Electronic Document

In order to talk about the development of a mathematical model for assessing alienability (or the risk of inalienability, which in this case is the same), it is necessary to formulate a problem statement.

Based on the general statement of the problem of ensuring long-term preservation, given in [5], the problem of assessing alienability can be formulated as follows.

Given:

1. The set of EDs $\boldsymbol{D} = \{ \boldsymbol{D_k} \}$
2. Requirements for allowable alienability values $\boldsymbol{\varphi_T}$
3. Mathematical model of long-term storage ED $\boldsymbol{D_k}$

Find:

1. Mathematical model for assessing alienability $\boldsymbol{\varphi(t)}$

The mathematical model of long-term storage ED is considered in detail in [4]. Let's bring it, because it will be important in solving the problem.

$$D_k = OrD \cup OdfD \cup DMD \cup CLI \cup LDI \quad (1)$$

OrD – this is an electronic document (original) or a digitized image of the original paper document, which we will also refer to as the original;

DMD – ED metadata, such as the author(s), time and place of creation, last modification time, format information, document name. Other data are possible, such as those set by the Dublin Core standard, see [6];

OdfD – normalized copy of ED. Normalization here refers to bringing the ED to a single format (set of formats) for long-term storage of the original document *OrD*;

CLI – this is reference information (classifiers, dictionaries, normative documents) to which the main document refers;

LDI – this is data about the documents associated with this ED.

3 Mathematical Model of the Alienability of an Electronic Document

The solution of the task should be a mathematical model for evaluating the alienation of the document. It is convenient to represent the alienation score as the probability that the ED can be completely removed from the firmware storage environment and transferred without loss to a new storage environment. Under the probability in this model we mean the assessment of the risk of non-alienation, performed, for example, with the help of an expert assessment. At the same time, risk levels should be set, for example, using a verbal scale of at least 3–5 positions (see, for example, [7]).

It is convenient to represent such an estimate as a weighted convolution of the probabilities of alienation of the semantic blocks of the long-term storage ED model (1).

Then the mathematical model of the probability of EL alienation can be represented as follows:

$$\varphi_k(D_k) = k_1 \cdot p\varphi OrD + k_2 \cdot p\varphi OdfD + k_3 \cdot p\varphi DMD + k_4 \cdot p\varphi CLI + k_5 \cdot p\varphi LDI \quad (2)$$

k_i – weight coefficients of the importance of the alienability of the semantic blocks of the ED (appointed by experts) $\sum k_i = 1, i = [1, 5]$. If there are no other preferences, then the author recommends assigning the following values $k_1 \geq 0{,}5$–$0{,}6$, $k_2 = \sum_{i=[3,5]} k_i$, because the alienability of the original ED (*OrD*) and the normalized copy (*OdfD*) is the most critical for long-term storage of ED.

$p\varphi OrD$ – the probability of alienation of the original (*OrD*) ED (including all its semantic blocks) from the storage environment;

$p\varphi OdfD$ – the probability of alienation of a normalized copy (*OdfD*) of an ED from the storage medium;

$p\varphi DMD$ – the probability of alienation of data on ED (*DMD*) from the storage environment;

$p\varphi CLI$ – the probability of alienation of reference information (*CLI*) related to the ED from the storage environment;

$p\varphi LDI$ – the probability of alienation of data on other documents related to ED (*LDI*) from the storage environment.

Of course, the presented mathematical model is valid for the evaluation of top-level semantic blocks. Each indicator of formula (2) must be detailed, as the mathematical model of ELD is detailed (see [4]).

4 Mathematical Model of the Alienability of the Original Electronic Document

The paper [4] provides a detailed mathematical model of the original ED, the composition of semantic blocks and their purpose. If we accept the assumptions presented in [4], the mathematical model for assessing the alienability of the original EDs of a typical electronic document management system (EDMS) is presented in the following form:

$$
\begin{aligned}
p\varphi OrD = {} & (\textstyle\sum_{(i=1,N1)} kN1_i (p\varphi OrDoc_i^{kN11} (\sum_{(j=1,M1)} kM1_{ij} p\varphi DSign_{ij})^{kN12}))^{k1} \\
& (\textstyle\sum_{(i=1,N2)} kN2_i (p\varphi OrRes_i^{kN21} (\sum_{(j=1,M2)} kM2_{ij} p\varphi RSign_{ij})^{kN22}))^{k2} \\
& (\textstyle\sum_{(i=1,N3)} kN3_i (p\varphi OrAgr_i^{kN31} (\sum_{(j=1,M3)} kM3_{ij} p\varphi ASign_{ij})^{kN32}))^{k3} \\
& (\textstyle\sum_{(i=1,N4)} kN4_i (p\varphi OrExe_i^{kN41} (\sum_{(j=1,M4)} kM4_{ij} p\varphi ESign_{ij})^{kN42}))^{k4} \\
& (\textstyle\sum_{(i=1,N5)} kN5_i (p\varphi OrMet_i^{kN51} (\sum_{(j=1,M5)} kM5_{ij} p\varphi MSign_{ij})^{kN52}))^{k5} \\
& (\textstyle\sum_{(i=1,N6)} kN6_i (p\varphi OrApp_i^{kN61} (\sum_{(j=1,M6)} kM6_{ij} p\varphi ApSign_{ij})^{kN62}))^{k6} p\varphi SignOrD^{k7}
\end{aligned}
\tag{3}
$$

where the weight coefficients of importance, assigned by experts or automatically, in the simplest case - equal to each other, must satisfy the conditions:

$$
\begin{aligned}
& \sum k_i = 1,\ i = [1,7], \\
& \sum_{(i=1,N1)} kN1_i = 1, \quad \sum_{(j=1,M1)} kM1_{ij} = 1, \quad kN11 + kN12 = 1, \\
& \sum_{(i=1,N2)} kN2_i = 1, \quad \sum_{(j=1,M2)} kM2_{ij} = 1, \quad kN21 + kN22 = 1, \\
& \sum_{(i=1,N3)} kN3_i = 1, \quad \sum_{(j=1,M3)} kM3_{ij} = 1, \quad kN31 + kN32 = 1, \\
& \sum_{(i=1,N4)} kN4_i = 1, \quad \sum_{(j=1,M4)} kM4_{ij} = 1, kN41 + kN42 = 1, \\
& \sum_{(i=1,N5)} kN5_i = 1, \quad \sum_{(j=1,M5)} kM5_{ij} = 1, kN51 + kN52 = 1, \\
& \sum_{(i=1,N6)} kN6_i = 1, \quad \sum_{(j=1,M6)} kM6_{ij} = 1, kN61 + kN62 = 1.
\end{aligned}
$$

$p\varphi OrDoc_i$ – the probability of alienation of the i-th semantic block ($OrDoc_i$) of the original ED (for example, if each page of a multipage document is represented by a separate digitized copy, several files make up one document (introduction, sections and conclusion), and other division options), each of which can be certified a separate set of cryptographic protection tools, for example, an electronic signature (ES) ($DSign_{ij}$);

$p\varphi OrRes_i$ – the probability of alienation of sheets of resolutions of the i-th original of the ED;

$p\varphi OrAgr_i$ – the probability of alienation of approval sheets of the i-th original of the ED;

$p\varphi OrExe_i$ – the probability of alienation of sheets of execution of the i-th original of the ED;

$p\varphi OrMet_i$ – the probability of alienation of familiarization sheets of the i-th original of the ED;

$p\varphi OrApp_i$ – the probability of alienation of the i-th ED application (for example, files, including audio and video, images, etc.);

$p\varphi DSign_{ij}$, $p\varphi RSign_{ij}$, $p\varphi ASign_{ij}$, $p\varphi ESign_{ij}$, $p\varphi MSign_{ij}$, $p\varphi ApSign_{ij}$ – probability of alienation of the j-th component of cryptoprotection (ES) of the i-th semantic block of the original ED ($OrDoc_i$, $OrRes_i$, $OrAgr_i$, $OrExe_i$, $OrMet_i$, $OrApp_i$);

$p\varphi SignOrD$ – the probability of alienation of cryptoprotection components (ES) that control the integrity of the original ED (OrD).

5 Mathematical Model of Alienability of a Normalized Copy of an Electronic Document

In addition to the original document (see (3)) it is also necessary to assess the alienability of other semantic blocks of the model (1). First of all, after the original ED, the alienability of the normalized copy is critical. This is due to the fact that the probability of interpreting a normalized copy after decades is higher than that of a document in the original format. Primarily because persistent storage formats are theoretically more interpretable. In [4], the model of a normalized copy of an ED is described in detail. According to this model, the assessment of alienability can be represented as follows:

$$p\varphi OdfD = \Big(\sum_{(i=1,N8)} kN8_i \left(p\varphi OdfDoc_i^{kN81} \left(\sum_{(il=1,L1)} kL1_{i1} p\varphi OdfLinkPic_{il} \right)^{kN82} \right)\Big)^{k8} \left(\sum_{(j=1,N9)} kN9_j p\varphi OdfPic_j \right)^{k9} \left(\sum_{(k=1,N10)} kN10_k p\varphi Sign_k \right)^{k10} \tag{4}$$

where the weight coefficients of importance, assigned by experts or automatically, in the simplest case - equal to each other, must satisfy the conditions:

$\sum k_i = 1$, $i = [8,10]$,

$\sum_{(i=1,N8)} kN8_i = 1$, $\sum_{(il=1,L1)} kL1_{i1} = 1$, $kN81 + kN82 = 1$,

$\sum_{(i=1,N9)} kN9_i = 1$, $\sum_{(i=1,N10)} kN10_i = 1$;

$p\varphi OdfDoc$ – the probability of alienation of the (normalized) text content of the semantic blocks [1–$N8$] of the original ED. See [8] for details on persistent storage formats. The semantic blocks of *OdfDoc* can contain sets of links to all graphic materials *OdfLinkPic*, the probability of their alienation is $p\varphi OdfLinkPic$ for links [1–$L1$];

$p\varphi OdfPic$ – the probability of alienating a set [1–$N9$] of normalized graphic information (raster and vector images, presentation elements, etc.) to be converted from source ED to graphic formats for long-term storage. See [8, 9] for details on formats;

$p\varphi Sign$ – the probability of alienation of a set of ES [1–$N10$] that certify a normalized document (contains signatory certificates, a chain of certificates, certification authority certificates, certificate revocation lists), see [4] for more details.

It is important that for a normalized document, you need to save both text materials in a persistent storage format, and the appearance in persistent image storage formats.

6 Development of a Mathematical Model of Metadata Alienability

The development of a mathematical model for the probability of metadata alienation $p\varphi DMD$ requires a separate study, which the author plans to conduct in the near future.

In the general case, the mathematical model will strongly depend on the composition of the ED metadata. In the simplest case, this will be a mathematical model of the alienability of 13 attributes in the main or 18 in the extended set of metadata of the standard [6].

For a complete model, it is necessary to add mathematical models of the alienability of the document content model, the visual form of the document presentation, various extracts from inventory logs (authenticity, interpretability), transaction logs, security logs, etc., as well as mathematical models of the alienability of indexes, including full-text ones.

7 Development of a Mathematical Model of Alienability of Related Data

Regulatory reference information (RRI), i.e. classifiers, dictionaries, normative documents (*CLI*), other EDs (*LDI*) referenced by the ED are also important for the interpretation of the ED in the future. If an ED that is subject to long-term storage refers to a specific RRI or other EDs, then these data must also be stored along with the EDs. Moreover, RRI and related EDs should be saved exactly in the version that was relevant at the time the ED was created. RRI and associated EDs may change over time, leading to the problem of incorrect interpretation of EDs.

Then the mathematical model of RRI alienability can be represented as follows:

$$p\varphi CLI = \sum_{(i=1,N11)} kN11_i \; p\varphi CLI_i \tag{5}$$

where $p\varphi CLI_i$ is the probability of alienation of the related RRI;

$kN11_i$ – the coefficient of importance of an individual element (semantic block) of the RRI, and $\sum_{(i=1,N11)} kN11_i = 1$. In the absence of a clear preference between the elements of the RRI, $kN11_i = (1/N11)$.

The main problem of using RRI is the automation of the classification of EDs. To solve this problem, there are several approaches: the first is to write the rules for classifying documents into classes, the second is to use machine learning.

In practice, the most reasonable is the combination of both approaches to solving the problem of automating the assignment of electronic documents to classes. See [10] for more details on the principles of constructing and training classifiers.

Similarly to the alienability of RRI, we can propose the following model of the alienability of related EDs:

$$p\varphi LDI = \sum_{(i=1,N12)} kN12_i \; p\varphi LDI_i \tag{6}$$

where $p\varphi LDI_i$ is the probability of alienation of the associated ED;

$kN12_i$ – coefficient of importance of a separate linked ED, and $\sum_{(i=1,N12)} kN12_i = 1$. In the absence of a clear preference between related EDs, $kN12_i = (1/N12)$.

8 An Example of a Scale of Values for Assessing the Probability of Non-alienation

A probabilistic model is used as a mathematical model for the alienation of EL. But, as mentioned above, the probability here is understood as the risk of non-alienation. How can you evaluate it.

If Bernoulli tests in the classical sense of probability assessment are impossible or time-consuming, then it is possible to use the verbal-numeric values of the risk assessment for the above indicators of the mathematical model. See [7] for details.

The assessment on this scale is performed by an expert (a group of experts), i.e. the person (persons) making the decision (DM), directly carrying out the organization of long-term storage of ED. Decision makers should be familiar with the hardware and software environment of long-term storage. An example of a value scale is shown in Table 1.

Table 1. An example of a verbal scale for a rough estimate of the likelihood of alienation.

Item number	Name of values	Numeric equivalent
1	Not alienated	0
2	Rather not alienate	0,25
3	Partially alienated	0,5
4	Rather alienated	0,75
5	Completely alienated	1

For a more accurate assessment of the probability of ED alienation, the software for long-term storage should include the function of periodic automatic inventory of the ED stock in order to automatically determine the possibility of alienating each ED from the hardware and software storage environment with automatic calculation of the probability of retrieval (Bernoulli test simulation).

If the alienation of the ED (or any semantic block of the ED) cannot be automatically performed, the personnel servicing the hardware and software storage environment must be notified of the violation of the alienation of a particular ED. It should also be possible to check the problematic ED for the possibility of its alienation in manual mode. If this attempt fails as well, the probability of retrieving this ED (or any semantic block of the ED) is automatically calculated by the ED storage system as equal to zero.

9 Conclusion

This article presents a mathematical modeling of the alienability of ED from the hardware and software storage environment. The presented mathematical model of ED alienation from the software and hardware storage environment is closely related to the mathematical model of long-term storage ED proposed by the author earlier (see [4]). The proposed mathematical model is probabilistic. However Bernoulli tests in the classical sense of probability assessment are impossible or time-consuming, the article shows how risk assessments can be used for the given indicators of the mathematical model.

The mathematical model is built taking into account the fact that the ED storage medium is always mobile and changeable, subject to parametric disturbances. Therefore, during long-term storage, the ED should be, as a control object, maximally stabilized in terms of the main characteristic - safety. To do this, in particular, it should be as independent as possible from the storage environment. As a measure of independence, a mathematical model for assessing the risk of inalienability of a document from the storage environment is proposed.

Mathematical models for evaluating the alienability of the semantic blocks of the original ED, the normalized copy of the ED, related RRI and other LD are proposed. In the future, it is planned to develop a mathematical model for assessing the alienability of ED metadata.

The calculations performed and the practical application of the approaches proposed in the article in a number of electronic archive projects, namely: electronic archives for the Pension Fund of the Russian Federation, Law Firm Gorodissky and Partners LLC, Cognitive Technologies LLC, give reason to believe that the presented model is adequate to the conditions of the long-term storage problem.

References

1. Solovyev, A.V.: The problem of defining the concept of “electronic document for long-term storage.” In: Silhavy, R., Silhavy, P., Prokopova, Z. (eds.) Data Science and Algorithms in Systems: Proceedings of 6th Computational Methods in Systems and Software 2022, Vol. 2, pp. 326–333. Springer International Publishing, Cham (2023). https://doi.org/10.1007/978-3-031-21438-7_26

2. Emelyanov, N.E.: Types of representation of structured data. Theoretical foundations of information technology. Collect. Works VNIISI **22**, 42–46 (1988)
3. Emelyanov, N.E.: Theoretical analysis of the document interface. All-Union Research Institute for System Research, Moscow, Russia (1987)
4. Solovyev, A.V.: Mathematical model of an electronic document of long-term storage. Inf. Technol. Comput. Syst. **2**, 30–36 (2022). https://doi.org/10.14357/20718632220204
5. Solovyev, A.V.: Long-term digital documents storage technology. In: Radionov, A.A., Karandaev, A.S. (eds.) Advances in Automation. LNEE, vol. 641, pp. 901–911. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-39225-3_97
6. GOST R 7.0.10–2019 (ISO 15836–1:2017) System of standards for information, librarianship and publishing. The Dublin Core metadata element set. Basic (core) elements (2019)
7. Petrovsky, A.B.: Decision theory. Akademy, Moscow, Russia (2009). ISBN: 978-5-7695-5093-5
8. Nikolayev, D.P., Postnikov, V.V., Usilin, S.A.: Cognitive PDF / A - technology for digitizing text documents for publication on the Internet and long-term archival storage. Proc. ISA RAS. **45**, 159–173 (2009)
9. Berestova, V.I.: Means and methods for creating an electronic document containing graphic images. J. Office Work. **1**, 45–56 (2014)
10. Smirnov, I., et al.: TITANIS: a tool for intelligent text analysis in social media. In: Kovalev, S.M., Kuznetsov, S.O., Panov, A.I. (eds.) Artificial Intelligence. LNCS (LNAI), vol. 12948, pp. 232–247. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-86855-0_16

Automatic Generation of an Algebraic Expression for a Boolean Function in the Basis {∧, ∨, ¬}

Roman Tsarev[1(✉)], Roman Kuzmich[2], Tatyana Anisimova[3], Biswaranjan Senapati[4], Oleg Ikonnikov[2,5], Viacheslav Shestakov[2], Alexander Pupkov[2], and Svetlana Kapustina[2]

[1] MIREA - Russian Technological University (RTU MIREA), Moscow, Russia
tsarev.sfu@mail.ru
[2] Siberian Federal University, Krasnoyarsk, Russia
[3] Kazan Federal University, Yelabuga, Russia
[4] University of Arkansas at Little Rock, Arkansas, IL, USA
[5] Reshetnev Siberian State University of Science and Technology, Krasnoyarsk, Russia

Abstract. Nowadays, there is a global trend of transition to new forms of education. E-learning platforms, in particular LMS Moodle, significantly expand the capabilities of the traditional form of education, give them new functions and provide opportunities that are not available in traditional education. Special mention can be made of the possibility of automated computer-based testing. Testing in general, is aimed at ensuring the objectivity of assessment of student knowledge. For the test to be effective, it is required that the instructor who composes the test has a high level of competence. Questions in the test should be varied and cover all the material studied. Particularly interesting are test questions that are unique to each individual student. As part of our study, we developed an algorithm that generates an algebraic expression for a Boolean function that the student needs to calculate. The answer obtained from the student is compared to the result of calculating the Boolean function, which is also performed automatically using Reverse Polish notation. #COMESYSO1120.

Keywords: E-learning · LMS Moodle · Automated Testing · Boolean Function · Formula Generation · Reverse Polish Notation · Algorithm

1 Introduction

The problem of organizing educational activities is one of the most significant in the field of education [1]. E-learning is extending or even replacing traditional forms of studying [2–7]. E-learning implies the use of information and communication technologies, interactive methods and techniques, technical means to optimize the learning process and improve students' knowledge [8–11]. A strong theoretical foundation obtained by students while studying specialized disciplines aimed at the development of professional skills will allow them to solve knowledge-intensive problems in various fields of science and technology in the future [12–14].

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 128–136, 2024.
https://doi.org/10.1007/978-3-031-53552-9_12

The specific features reflecting the advantages of e-learning include:

- interactivity;
- multimedia;
- joint activities not tied to place and time (use of chat and/or video conferencing),
- development of networked virtual mobility;
- use of a variety of active learning methods;
- individualization and differentiation of the learning process;
- increased opportunities of access to electronic educational resources;
- improved efficiency of organization of students' and teacher's time by automating the performance of routine tasks;
- digital forms and types of control, automation of results evaluation.

Currently, there is a large number of e-learning platforms [15–19]. One of the most popular is LMS Moodle (Learning Management System Modular Object-Oriented Dynamic Learning Environment). LMS Moodle provides the existence and functioning of electronic courses in a software environment, remote interaction between teachers and students, including to support face-to-face learning [20–23]. The modular structure of Moodle and advanced functionality for e-learning purposes allow to create and store e-learning materials, setting the sequence of their studying and monitoring the knowledge, skills and abilities of students within the e-course created by the teacher [24–26]. LMS Moodle allows the teacher to control the time of students' work in the system as a whole, and in individual modules of the course. The automatic assessment of the students' work results by the LMS Moodle system allows the teacher to view the marks on the tests, to analyze the students' progress due to the acquisition of statistics. Thus, LMS Moodle is characterized by the following features: interactivity, flexibility, scalability and standardization.

Along with such components of e-course within LMS Moodle as video lectures and lecture-presentations, workshops, seminars, practical and laboratory classes, projects and online conferences, forum, blog, etc., there is a knowledge control block, which deserves special attention. Such a module with control-measuring materials (test tasks) can be used for a current, intermediate or final control [1, 27].

The content of the test tasks must fully comply with the content of the course, provided by the state standard and the curriculum [1]. The content of the test tasks must fully comply with the content of the discipline, provided by the state standard and the curriculum. In addition to test tasks there can be used active control methods: discussions, problem, game and simulation situations, web quests, case studies, interactive quizzes, etc.

Automated testing technology (computer-based testing) - is an innovative form of knowledge control, which allows to instantly and automatically check and evaluate the results, to obtain an objective and independent characteristics of the level of educational achievement. The following advantages of automated testing can be noted:

- automatic testing of knowledge;
- operative, automatic, objective diagnostics of results;
- possibility of generating a large number of test variants;
- convenient procedure for entering and modifying test materials;

- possibility to manage test content, difficulty levels, test strategies;
- time-saving for students and teachers.

The technique of creating a computer-based test is closely connected with the choice of test form, with the goal and learning objectives set, the content of the test itself, the level of student preparedness. These aspects should be taken into account at the stage of test creation. In general, there are a number of requirements that a teacher needs to observe when planning a computerized test:

- clear, explicit, concise text of the tasks;
- a set of test questions should cover all the learning material necessary for mastering;
- in order to exclude the mechanical memorization of test sequence and "copying", test questions should fall in random order, answer options should also be generated randomly, in random order;
- ability to set the number of test attempts and time delays between attempts;
- possibility to customize the mode of viewing the results.

The main quality criteria of an automated test are efficiency, reliability, differentiation, security, multimedia, completeness of test item types, ease of use.

Efficiency refers to completeness of the test, comprehensiveness of testing, proportionality of the presentation of all elements of the studied items of knowledge.

Reliability of computer testing is characterized by stability, sustainability of indicators for repeated measurements with the same test or its equivalent substitute.

Differentiated test is called such testing, through which it is possible to divide students into those who have mastered the material at the required level and those who have not reached this level.

Multimedia shows in the possibility to embed multimedia objects (sound, video, graphics) in the test tasks.

Security is to ensure that the test takers cannot access the reference answers, and students and outsiders to the overall test results.

Completeness of types of test tasks implies mandatory support for five basic types of test tasks (open-type and closed-type: single and multiple choice, matching, sequencing).

Ease of use for teacher and student implies a standard modern interface, clear and effective division into functional subsystems, focus on standards in the information industry, a common style of software design, documentation for each user group, transparent meaning of configurable parameters, clear and easy for teacher to present test results (including assessment of test quality).

Certainly, the above criteria are partly idealized, not all automated tests include these requirements, this is something that every teacher should strive for to improve the effectiveness of the educational process.

Our research concerns the automatic generation of a test question on the course of mathematical logic. Test question involves the calculation of a Boolean function by the student. To ensure that each student was offered its own unique version of a test question an algorithm for generation of an algebraic expression for a Boolean function was designed and implemented. In addition to this Boolean function is also calculated automatically using Reverse Polish notation. This allows to compare the true result with the student's answer and assign the appropriate grade. This fully automates the process

of creating a question, answering it, comparing it to the student's answer, and grading the student's answer. The algorithm of generation of an algebraic expression for a Boolean function is presented in the next section. The proposed approach and algorithm can also be used in the discrete mathematics course, as well as in other courses in which Boolean functions are studied.

2 Method

Testing in LMS Moodle is automatic and does not require a teacher [28–30]. The student gets a test question and enters the answer to it. In this study, each student gets a different Boolean function to solve it. In order to generate Boolean functions while students are testing, the algorithm of generation for an algebraic expression for a Boolean function was developed. Generating a Boolean function, operators (conjunction, disjunction and negation) and variables are chosen randomly. Thus, the algebraic expression is generated in the basis of Boolean functions $\{\wedge, \vee, \neg\}$.

When answering a test question, the student has to make a truth table and enter the answer into Moodle. In order to check the student's answer, the generating Boolean function is calculated using Reverse Polish notation. After that the student's answer and the result of Boolean function calculation are compared and the number of points the student got for this question is calculated. When generating a Boolean function, a tree is built with variables and operators of conjunction, disjunction, and negation as its nodes.

Algorithm for creating a tree representing an algebraic expression for a Boolean function

Step 1. Create a node without label[a)]
Step 2. If the current node is unlabeled, assign it a label chosen at random[b)]:
- variable x_i ($i = 1...n$, n is the number of variables)[c)]; go to step 3
- unary operator (negation $\neg$); go to step 4
- binary operator (conjunction $\wedge$ or disjunction $\vee$); go to step 5

otherwise, if the node has a label, then:
if it is a root node, the algorithm stops;
otherwise, if the node is labeled with unary operator ($\neg$),
then move to the parent of this node; go to step 2;
otherwise ($\wedge$ or $\vee$), if the node has no right child,
then create the right edge; go to step 1;
otherwise, go to the parent of this node;
go to step 2;

Step 3. If the current node is the root of the tree then algorithm stops;
otherwise, go to the parent of the current node; go to step 2;
Step 4. Create an edge to the next level; go to step 1[d)];
Step 5. Create the left edge; go to step 1.

Note:

a) At the first step of the algorithm the tree depth l can be set. In this case, if the current node is at level l, the node has to be marked with variable x_i.

b) The random number generator can be supplemented with weights which will determine the probability of a variable, unary or binary operator appearing in the tree.
c) Variables can be taken sequentially or randomly. In the second case, when the algorithm stops, it will be necessary to check the tree leaves to make sure that all variables are present in their labels. If this is not the case, then replace the duplicated variables with the missing variables.
d) If double negation in the generated algebraic expression is to be excluded, then step 4 has to be rewritten as follows:

Step 4. If the parent node is labeled "negation", then remove the label from the current node; go to step 1.

3 Results and Discussion

Figure 1 shows the tree generated using the algorithm for creating a tree representing an algebraic expression for a Boolean function.

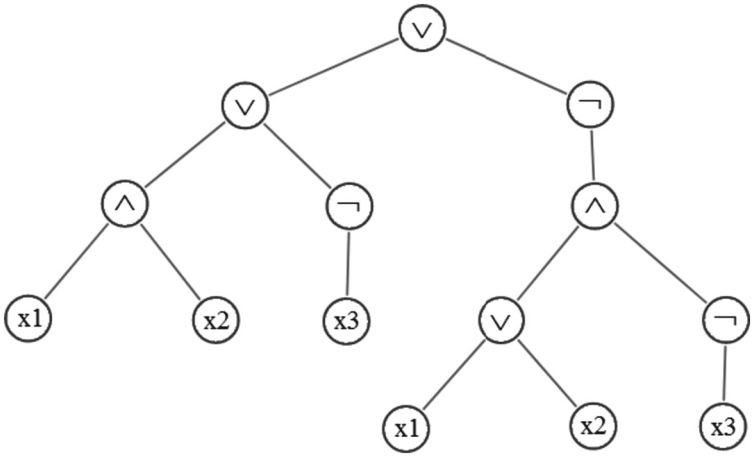


Fig. 1. A Boolean function represented as a tree.

The generated algebraic expression for a Boolean function shown in Fig. 1 is:

$$x_1 \wedge x_2 \vee \neg x_3 \vee \neg(x_1 \vee x_2 \wedge \neg x_3). \tag{1}$$

The advantage of this tree is that it can be used to calculate function values using Reverse Polish notation. By traversing the tree starting from the bottom left node according to Reverse Polish notation application we obtain for the generated Boolean function the following algebraic expression:

$$x_1 x_2 \wedge x_3 \neg \vee x_1 x_2 \vee x_3 \neg \wedge \neg \vee .$$

To calculate a Boolean function using Reverse Polish notation, a complete brute-force over all sets of variable values is done. In this case, when there are three variables, these are the sets 000, 001,…, 111.

As a result, we get the truth table shown in Table 1.

Table 1. Table of truth for the Boolean function (1).

x_1	x_2	x_3	$f(x_1, x_2, x_3)$
0	0	0	1
0	0	1	1
0	1	0	1
0	1	1	1
1	0	0	1
1	0	1	0
1	1	0	1
1	1	1	1

The result of calculating the Boolean function (1) can be represented as a vector function $f(x_1, x_2, x_3) = (11111011)$.

When performing the test, the student has to calculate a truth table by calculating the values of the given Boolean function.

The student's answer entered into the LMS Moodle testing module is compared to the calculated values of the function. The grade for this question is calculated as follows:

the number of correct Boolean function values provided by the student/2^n,

where n is the number of variables in the Boolean function.

So, if the student makes two mistakes when calculating function values from three variables, he/she will receive a grade of 0.75 for this question.

4 Conclusion

The trend toward informatization of education and the active use of information and communication technologies everywhere has led us to e-learning. In both full-time and part-time education e-learning systems help organize independent work and continuous monitoring of the educational process.

Due to other specific features of e-learning (interactivity, multimedia, individualization and differentiation of learning) new opportunities of interaction are opening up for the teacher and students.

One of the main tools of such interaction is an electronic educational platform. LMS Moodle has a set of features that allow to change easily the learning process depending on the curriculum of the higher education institution, as well as methods and means of learning. LMS Moodle is a standard of distance and blended learning, a means of

improving the quality of education. This system allows a large number of teachers and students to work simultaneously, to create electronic courses.

Conducting automated testing, as part of an e-course, is a unique opportunity for objective knowledge control. Automated tests are a standardized system of tasks, allowing a reliable and objective assessment of the level of students' knowledge, skills and abilities. Testing can be carried out without participation of the teacher, thus students develop such personal qualities as self-control and self-organization. The computer-based test also has the possibility of including multimedia objects, expanding the range of responses of the test taker. The main advantage is the ability to automate mathematical and statistical processing of results. In order to get a high-quality, effective test, it is important to comply with the requirements for its design and, most importantly, for the development of tasks and answer options.

With the help of the algorithm for creating a tree representing an algebraic expression for a Boolean function, we managed to generate variants of test tasks so that each student has his own variant. This method helps to avoid borrowing answers between students with a huge number of generated individual variants, increases the objectivity of the obtained results, contributes to the individualization of learning.

An important task of learning is the formation of students' abilities to solve professional problems. The possibilities of e-learning, educational platform LMS Moodle and automated testing make this process effective, increase motivation for learning, form the information culture and media competence of the teacher and students, increase the visibility and virtuality of learning.

References

1. Gushchin, A.N.: Design-based science: curriculum for architects. In: AIP Conference Proceedings, vol. 2657, p. 020047 (2022)
2. Al-Fraihat, D., Joy, M., Masa'deh, R., Sinclair, J.: Evaluating e-learning systems success: an empirical study. Comput. Hum. Behav. **102**, 67–86 (2020). https://doi.org/10.1016/j.chb.2019.08.004
3. Deetjen-Ruiz, R., et al.: Applying ant colony optimisation when choosing an individual learning trajectory. In: Silhavy, R., Silhavy, P. (eds.) Networks and Systems in Cybernetics. CSOC 2023. Lecture Notes in Networks and Systems, vol. 723, pp. 587–594 (2023). https://doi.org/10.1007/978-3-031-35317-8_53
4. Lin, H.-M., et al.: A review of using multilevel modeling in e-learning research. Comput. Educ. **198**, 104762 (2023). https://doi.org/10.1016/j.compedu.2023.104762
5. Aljarbouh, A., Ahmed, M.S., Vaquera, M., Dirting, B.D.: Intellectualization of information processing systems for monitoring complex objects and systems. Modern Innov. Syst. Technol. **2**(1), 9–17 (2022). https://doi.org/10.47813/2782-2818-2022-2-1-9-17
6. Lunev, D., Poletykin, S., Kudryavtsev, D.O.: Brain-computer interfaces: technology overview and modern solutions. Modern Innov. Syst. Technol. **2**(3), 0117–0126 (2022). https://doi.org/10.47813/2782-2818-2022-2-3-0117-0126
7. Maulana, F.I., Febriantono, M.A., Raharja, Khaeruddin, Herasmara, R.: Twenty years of e-learning in health science: a bibliometric. Procedia Comput. Sci. **216**, 604–612 (62023). https://doi.org/10.1016/j.procs.2022.12.175
8. Al-smadi, A.M., Abugabah, A., Al Smadi, A.: Evaluation of e-learning experience in the light of the covid-19 in higher education. Procedia Comput. Sci. **201**, 383–389 (2022). https://doi.org/10.1016/j.procs.2022.03.051

9. Kacetl, J., Semradova, I.: Reflection on blended learning and e-learning – case study. Procedia Comput. Sci. **176**, 1322–1327 (2020). https://doi.org/10.1016/j.procs.2020.09.141
10. Singh, S., Singh, U.S., Nermend, M.: Decision analysis of e-learning in bridging digital divide for education dissemination. Procedia Comput. Sci. **207**, 1970–1980 (2022). https://doi.org/10.1016/j.procs.2022.09.256
11. Akhmetjanov, M., Ruziev, R.: Fundamentals of modeling fire safety education. Inform. Econ. Manag. **1**(2), 0301–0308 (2022). https://doi.org/10.47813/2782-5280-2022-1-2-0301-0308
12. Senapati, B., Rawal, B.S.: Adopting a deep learning split-protocol based predictive maintenance management system for industrial manufacturing operations. In: Hsu, CH., Xu, M., Cao, H., Baghban, H., Shawkat Ali, A.B.M. (eds.) Big Data Intelligence and Computing. DataCom 2022. Lecture Notes in Computer Science, vol. 13864, pp. 22–39. Springer, Singapore (2023). https://doi.org/10.1007/978-981-99-2233-8_2
13. Naeem, A.B., Senapati, B., Islam Sudman, M.S., Bashir, K., Ahmed, A.E.M.: Intelligent road management system for autonomous, non-autonomous, and VIP vehicles. World Electr. Veh. J. **14**(9), 238 (2023). https://doi.org/10.3390/wevj14090238
14. Sabugaa, M., Senapati, B., Kupriyanov, Y., Danilova, Y., Irgasheva, S., Potekhina, E.: Evaluation of the prognostic significance and accuracy of screening tests for alcohol dependence based on the results of building a multilayer perceptron. In: Silhavy, R., Silhavy, P. (eds.) Artificial Intelligence Application in Networks and Systems. CSOC 2023. Lecture Notes in Networks and Systems, vol. 724, pp. 240–245. Springer, Cham (2023).https://doi.org/10.1007/978-3-031-35314-7_23
15. Nácher, M.J., Badenes-Ribera, L., Torrijos, C., Ballesteros, M.A., Cebadera, E.: The effectiveness of the GoKoan e-learning platform in improving university students' academic performance. Stud. Educ. Eval. **70**, 101026 (2021). https://doi.org/10.1016/j.stueduc.2021.101026
16. Natasia, S.R., Wiranti, Y.T., Parastika, A.: Acceptance analysis of NUADU as e-learning platform using the technology acceptance model (TAM) approach. Procedia Comput. Sci. **197**, 512–520 (2022). https://doi.org/10.1016/j.procs.2021.12.168
17. Rujuan, W., Lei, W.: Research on e-learning behavior evaluation of students based on three-way decisions classification algorithm. Procedia Comput. Sci. **208**, 367–373 (2022). https://doi.org/10.1016/j.procs.2022.10.052
18. Veeramanickam, M.R.M., Ramesh, P.: Analysis on quality of learning in e-Learning platforms. Adv. Eng. Softw. **172**, 103168 (2022). https://doi.org/10.1016/j.advengsoft.2022.103168
19. Zybin, D., Kalach, A., Rogonova, A., Bashkatov, A., Klementeva, M. Structural and parametric synthesis of a document management system. Modern Innov. Syst. Technol. **1**(4), 24–30 (2021). https://doi.org/10.47813/2782-2818-2021-1-4-24-30
20. Benguedach, A., Boudia, C., Bouamrane, K.: Interpretive analysis of online teaching labs constructed using moodle during the pandemic period. Heliyon **9**(5), e16007 (2023). https://doi.org/10.1016/j.heliyon.2023.e16007
21. De Medio, C., Limongelli, C., Sciarrone, F., Temperini, M.: MoodleREC: a recommendation system for creating courses using the moodle e-learning platform. Comput. Hum. Behav. **104**, 106168 (2020). https://doi.org/10.1016/j.chb.2019.106168
22. Rezgui, K., Mhiri, H., Ghédira, K.: Extending moodle functionalities with ontology-based competency management. Procedia Comput. Sci. **35**, 570–579 (2014). https://doi.org/10.1016/j.procs.2014.08.138
23. Tsarev, R., et al.: Gamification of the graph theory course. finding the shortest path by a greedy algorithm. In: Silhavy, R., Silhavy, P. (eds.) Networks and Systems in Cybernetics. CSOC 2023. Lecture Notes in Networks and Systems, vol. 723, pp. 209–216. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-35317-8_18

24. Dascalu, M.-D., et al.: Before and during COVID-19: a cohesion network analysis of students' online participation in moodle courses. Comput. Hum. Behav. **121**, 106780 (2021). https://doi.org/10.1016/j.chb.2021.106780
25. Tsarev, R.Y., et al.: An approach to developing adaptive electronic educational course. Adv. Intell. Syst. Comput. **986**, 332–341 (2019). https://doi.org/10.1007/978-3-030-19813-8_34
26. Yamaguchi, S., Kondo, H., Ohnishi, Y., Nishino, K.: Design of question-and-answer interface using moodle Database function. Procedia Comput. Sci. **207**, 976–986 (2022). https://doi.org/10.1016/j.procs.2022.09.153
27. Gushchin, A.N., Divakova, M.N.: Nurturing cognitive skills in undergraduates with the help of ontological analysis. In: IOP Conference Series: Materials Science and Engineering, vol. 972, p. 012059 (2020). https://doi.org/10.1088/1757-899X/972/1/012059
28. Tsarev, R., et al.: Improving test quality in e-learning systems. In: Silhavy, R., Silhavy, P. (eds.) Networks and Systems in Cybernetics. CSOC 2023. Lecture Notes in Networks and Systems, vol. 723, pp. 62–68 (2023). Springer, Chamhttps://doi.org/10.1007/978-3-031-35317-8_6
29. Shilova, T.V., Artamonova, L.V., Averina, S.: Computer-based tests as an integral component of an EFL course in moodle for non-linguistic students. Procedia Soc. Behav. Sci. **154**, 434–436 (2014). https://doi.org/10.1016/j.sbspro.2014.10.187
30. Yurchenko, P.: Ways to solve the problem of documentary thematic search. Inform. Econ. Manag. **2**(1), 0101–0123 (2023). https://doi.org/10.47813/2782-5280-2023-2-1-0101-0123

Applying a Recurrent Neural Network to Implement a Self-organizing Electronic Educational Course

Ruslan Khakimzyanov[1], Sadaquat Ali[2], Bekbosin Kalmuratov[3], Phuong Nguyen Hoang[4], Andrey Karnaukhov[5], and Roman Tsarev[6,7(✉)]

[1] Tashkent State Transport University, Tashkent, Uzbekistan
[2] University of Warwick, Coventry, UK
[3] The Nukus branch of the Tashkent University of Information Technologies, Tashkent, Uzbekistan
[4] University of Social Sciences and Humanities, Vietnam National University, Ho Chi Minh City (VNU-HCM), Vietnam
[5] Reshetnev Siberian State University of Science and Technology, Krasnoyarsk, Russia
[6] MIREA - Russian Technological University (RTU MIREA), Moscow, Russia
tsarev.sfu@mail.ru
[7] Bauman Moscow State Technical University, Moscow, Russia

Abstract. This article discusses the features of the implementation of a self-organizing electronic educational course as an innovative form of learning. A recurrent neural network was used to implement the self-organizing e-course. A recurrent neural network has the ability to remember previous data and use it to make decisions in the current iteration. This feature allows to take into account the current level of student knowledge and based on it provide test questions. The practical part of the study presents the results of testing the self-organizing electronic course created using recurrent neural network on a group of students. The results confirm the increase of learning efficiency and improvement of students' results #COMESYSO1120.

Keywords: E-learning · E-course · Self-organizing Course · Moodle · Recurrent Neural Network · Elman Network

1 Introduction

E-learning since its first appearance has quickly established itself as an effective and powerful approach to learning [1–4]. Its advantages over traditional forms of education are so significant and fully meet the needs of modern learners, that its popularity and widespread use in practice is inevitable [5–8].

Today e-learning is applied not only in education, but also in other spheres, including business, public authorities and other spheres of professional activity [9–14]. Companies use e-learning to train their employees, and educational organizations use it to conduct

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 137–145, 2024.
https://doi.org/10.1007/978-3-031-53552-9_13

online courses and seminars [15–18]. E-learning is becoming more and more popular among people seeking education nowadays, where mobility and flexibility are an important part of everyday life.

The rationale for the widespread practical use of e-learning is its main advantages such as:

- Flexibility. E-learning allows students to learn at any time and at any place.
- Interactivity. Education through e-learning includes many interactive elements, such as exercises, videos, audio, tests, etc.;
- Personalization. E-learning platforms can be customized to the individual needs of the learner and also allow selection of training programs in accordance with the market demand.
- Cost-effectiveness. Educational materials for learning through online courses are always much cheaper than for learning in traditional educational institutions.
- Mobility. E-learning can be accessed not only on a PC or laptop, but also through mobile devices such as phones and tablets;
- Efficiency. E-learning shows high efficiency because it allows students to access information and evaluate their progress in real time;
- Adaptability. E-learning responds immediately to changes in the industry and updates educational and training programs to meet new requirements.

In today's world, e-learning is becoming increasingly popular, and there are a number of reasons why this approach will continue to evolve in the future.

The growth of the e-learning industry. The global environment during the last few years, due to the COVID-19 pandemic, has led to a surge in demand for online education and the e-learning technology market. In the future, it is clear that e-learning will continue to grow in popularity, with more countries and individual learners recognizing the benefits of this approach [19].

Development of new technologies. The capabilities of such technologies will continue to contribute to the development of e-learning. Improved virtual and augmented realities, blockchain and the development of artificial intelligence will subsequently also help improve the possibilities for interactivity and personalization of materials.

Adaptability of learning. Today's technology makes it possible to quickly and successfully create a variety of adaptive courses that can adjust to the needs of learners. This helps them achieve better results and significantly accelerate the learning process.

Lifelong Learning concept. Lifelong learning is becoming increasingly important and relevant these days, and e-learning is certainly an ideal tool for this approach. Courses can be targeted at different age groups and levels of knowledge, which will help support the continuous development and updating of skills and in practice is extremely user-friendly.

Globalization of learning. E-learning also allows learners from different countries to receive education from the best educators in the world and to attend "electronic" universities. Thanks to this, students have the opportunity to study all over the world, broaden their horizons, and interact more closely with students and teachers from different cultural contexts. Territorial boundaries for higher education are no longer the limit, and this trend is only going to grow in the future.

Convenience, accessibility, and ease of use. E-learning provides learners with the opportunity to study anytime and anywhere using their computer, other Internet-connected devices or cell phones, which makes learning more convenient and accessible for many people.

The development of e-learning, its increasing spread in practice eventually leads to the emergence of qualitatively new forms and methods of learning. In this direction, of course, e-learning will also progressively develop [20]. For example, the main modern methods of e-learning include:

1. Virtual classes are online courses that are created by instructors using a variety of personalized tools and learning materials.
2. Mobile learning is the use of mobile devices for learning, for learning theoretical material, additional information, and for assignments. Mobile devices provide access to resources via the Internet and provide learners with the ability to learn anytime and anywhere.
3. Video lessons are recorded lessons that can be viewed online or downloaded to a computer. The instructor speaks into the camera, and students can view or download the recordings at any time.
4. Learning management systems are software that allows to organize, monitor, and evaluate learning. It includes various tools for creating learning materials, testing, assessment, and feedback.

In general, the use of e-learning allows students to gain knowledge and skills more effectively and conveniently, opens up access to a wide range of learning materials and saves time.

At present, the so-called self-organizing course is becoming more and more relevant, which allows to increase the efficiency of learning and is as much as possible focused on the needs of a particular student [21–23].

A self-organizing e-course is an online course that uses methods and technology to adjust the course and adapt it to the needs of students.

Such courses are based on a methodology that allows the course material to be adapted to the changing needs and interests of students and their learning process. The theoretical materials and assignments for each lesson are already defined, but changes can be done according to the intermediate students results reflecting their level of knowledge.

The self-organizing e-course also uses analytical tools to examine student behavior to determine the most effective approaches to learning. This improves the quality of the course and enhances learning.

Overall, self-organizing e-course is a promising direction in online education that provides flexibility in the learning process and better quality of learning based on high-tech methods and analytics.

2 Method

The main characteristic of a self-organizing electronic course is its division into modules. Each module includes theoretical and practical material, grouped by topics, at the end of the module the student has to pass a test on the studied material.

Note that the student does not have to pass the test qualitatively the first time - if he or she fails it, then in the future, when passing the next module, incorrectly solved questions will be presented to the students for solving again. In turn, when the student moves on to a new test, the questions from past topics will reappear in front of him. All this leads to the formation of a kind of self-configured learning course with effective feedback.

Artificial intelligence and, in particular, neural networks are actively developing today [24–26]. In our work we used a recurrent neural network to implement the self-organizing e-course [27]. A recurrent neural network is ideal for creating such a course because it has the ability to remember previous data and use it to make decisions at the current iteration. That is, these neural networks have a memory effect. Recurrent neural networks allow to offer tasks of a certain level of difficulty depending on the learner's behavior, and they are also used to implement models that determine which materials a student should re-learn when the test score for a particular module is low.

It is noteworthy that recurrent neural network allows to implement adaptive testing, in which tests are optimized according to the student's previous answers. In particular, if a student showed high performance and answered most of the questions correctly, he can be offered more difficult test tasks. At the same time, in case of low test scores the student may be offered additional study materials for successful continuation of the course.

In this study an Elman network [28–31] was used (see Fig. 1). The Elman network is a three-layer perceptron including three layers: input, output and hidden. However, besides this, there is also a context layer, which makes it possible to take into account the prehistory of answers to test questions and to accumulate information for further testing in the optimal way for a given student.

The equations describing the internal state and output of the Elman network are as follows:

$$h^{(t)} = f(Ux^{(t)} + Wh^{(t-1)} + b),$$

$$\tilde{y}^{(t)} = g(Vh^{(t)} + c),$$

where U, W, V are matrices of weights;

b, c are shift vectors;

$h^{(t)}$ is the vector of hidden variables at the moment t;

$\tilde{y}^{(t)}$ is the output of the Elman network at the moment of t;

$f(\cdot)$, $g(\cdot)$ are activation functions.

The major goal of training the Elman network is to reduce the cumulative error:

$$E = \frac{\sum_{i=1}^{N} \sum_{t=1}^{\tau_i} d(y_i^{(t)}, \tilde{y}_i^{(t)})}{N} \to \min_{U,W,V},$$

E indicates the difference between the real output and the output of the Elman network. N denotes the number of input sequences of the Elman network; the elements of the i-th sequence are denoted by τ_i; $y^{(t)}$ and $\tilde{y}^{(t)}$ are the real output and the output

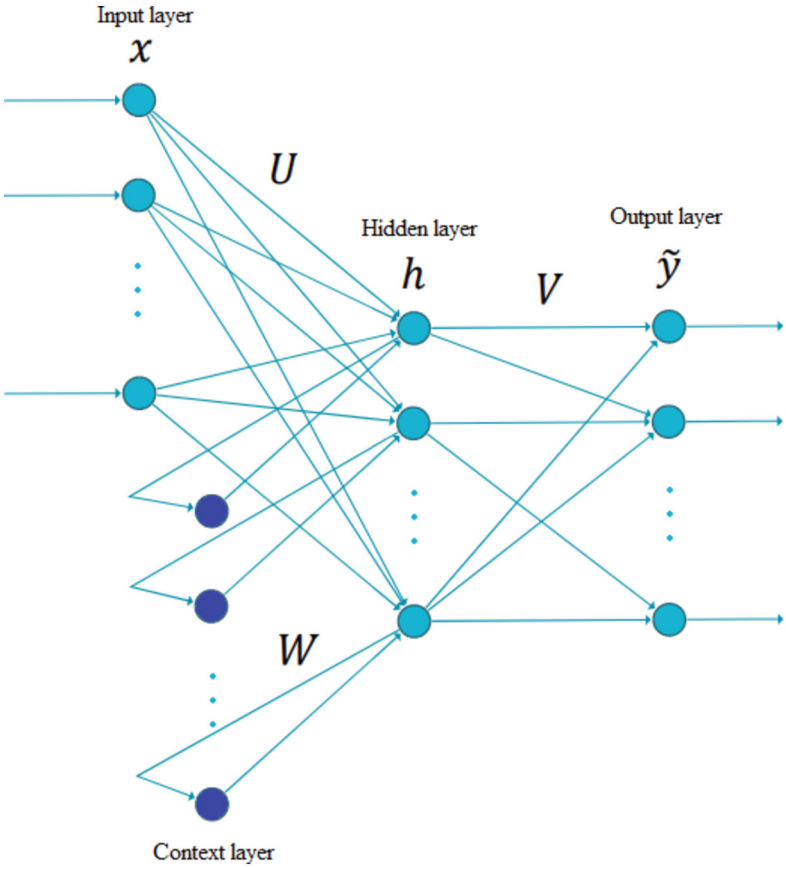


Fig. 1. The Elman network.

of the Elman network at time t at the i-th input sequence, respectively. The Euclidean distance or cross-entropy can be used to measure the difference d between the real output and the Elman network output. It is required to find such values of parameters U, W, V that the value of E is minimal.

The Elman network training used backpropagation through time, positively proven in solving many practical problems [32–36].

Implementation of self-organizing e-course became possible through the use of LMS Moodle, which allows to maximize the potential of e-learning. In particular, LMS Moodle provides an opportunity of digital setting of e-course in such a way that while passing the next module, the student is also provided with topics from the previous module (or modules), which he learned insufficiently well, that is, received a low score when answering questions on this topic during testing. The number of scores in the test is also stored in LMS Moodle and linked directly to the digital profile of the student.

3 Results and Discussion

For the practical part of the study we took eight groups of students with a total number from 26 to 32 people. The test subjects were divided into 2 blocks of 4 groups each. In the first, students were taught using the improved, described above methodology using a self-organizing course. In turn, in the second block, students continued to study in the traditional format without the use of new technologies.

Let us briefly outline the general characteristics of the course used in the study: it consists of 8 modules, the final test includes questions selected at random from the bank of questions, which was also used in the current testing during the course.

Features of the mechanism of self-organizing course contributed to the fact that the students of the 1st block throughout the course periodically repeated the material learned earlier not well, which subsequently helped to more objectively assess the quality of the final test of the course.

According to the results of the study we obtained the following results. In the case of students in block 1, working on an innovative method, the number of points scored for the test was consistently higher than in the case of students in block 2, studying traditionally. The difference in the arithmetic mean of the test scores of block 1 and block 2 students ultimately ranged from 8% to 27%.

The obtained results allow to conclude that the teaching method proposed in this work is really effective in use.

In addition, the above final data show that the self-organizing electronic course allows directly in the learning process to identify gaps in the knowledge of students, to determine the unlearned material and through repeated study promptly and qualitatively eliminate gaps in learning.

4 Conclusion

E-learning is a modern approach to learning, based on the use of information technology and new learning methods. E-learning involves learning on computers or mobile devices, using virtual classroom technology, multimedia, interactive courses, webinars, chat rooms, forums and other formats.

E-learning allows studying independently, in a comfortable place and time pace, and can be carried out both remotely and offline. E-learning is used in various fields, including education, business, medicine, and public administration. It provides students with the opportunity to gain qualifications, acquire new knowledge and skills, and additional knowledge in their professional field.

E-learning has great potential for effective and convenient learning. Its main advantages are related to the additional opportunities it gives to students and teachers, as well as saving time and resources.

The self-organized e-course is one of the most relevant nowadays innovative forms of learning, which seems to be much more effective and convenient for students and teachers. It is promoted, in particular, by the special organization of an electronic course which assumes a modular format with an opportunity to repeatedly repeat the passed material through the solution of questions which have caused the greatest difficulty to the student.

To implement a self-organizing e-course a recurrent neural network - Elman network was used, which allowed to take into account the results of studying the material by the student and make the testing more relevant to his current level.

In order, to confirm the practical significance of the presented learning techniques, we tested it on several groups of students. The results of this study, presented in the work, confirmed the effectiveness of this approach, its high relevance for use in the practical activities of teachers.

References

1. Alyoussef, I.Y.: Acceptance of e-learning in higher education: the role of task-technology fit with the information systems success model. Heliyon **9**(3), e13751 (2023). https://doi.org/10.1016/j.heliyon.2023.e13751
2. Balogun, N.A., Adeleke, F.A., Abdulrahaman, M.D., Shehu, Y.I., Adedoyin, A.: Undergraduate students' perception on e-learning systems during COVID-19 pandemic in Nigeria. Heliyon **9**(3), e14549 (2023). https://doi.org/10.1016/j.heliyon.2023.e14549
3. Behl, A., Jayawardena, N., Pereira, V., Islam, N., Del Giudice, M., Choudrie, J.: Gamification and e-learning for young learners: a systematic literature review, bibliometric analysis, and future research agenda. Technol. Forecast. Soc. Chang. **176**, 121445 (2022). https://doi.org/10.1016/j.techfore.2021.121445
4. Ouajdouni, A., Chafik, K., Boubker, O.: Measuring e-learning systems success: data from students of higher education institutions in Morocco. Data Brief **35**, 106807 (2021). https://doi.org/10.1016/j.dib.2021.106807
5. Chahal, J., Rani, N.: Exploring the acceptance for e-learning among higher education students in India: combining technology acceptance model with external variables. J. Comput. High. Educ. **34**, 844–867 (2022)
6. Hsu, H.-P., Guo, J.-L., Lin, F.-H., Chen, S.-F., Chuang, C.-P., Huang, C.-M.: Effect of involvement and motivation on self-learning: Evaluating a mobile e-learning program for nurses caring for women with gynecologic cancer. Nurse Educ. Pract. **67**, 103558 (2023). https://doi.org/10.1016/j.nepr.2023.103558
7. Sayaf, A.M.: Adoption of E-learning systems: an integration of ISSM and constructivism theories in higher education. Heliyon **9**(2), e13014 (2023). https://doi.org/10.1016/j.heliyon.2023.e13014
8. Tsarev, R., et al.: Improving test quality in e-learning systems. In: Silhavy, R., Silhavy, P. (eds.) Networks and Systems in Cybernetics. CSOC 2023. Lecture Notes in Networks and Systems, vol. 723, pp. 62–68. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-35317-8_6
9. Baranov, I.: Review and comparative analysis of BPMN-systems for robotization of business processes. Modern Innov. Syst. Technol. **2**(3), 0139–0149 (2022). https://doi.org/10.47813/2782-2818-2022-2-3-0139-0149
10. Bloomfield, J.G., Fisher, M., Davies, C., Randall, S., Gordon, C.J.: Registered nurses' attitudes towards e-learning and technology in healthcare: a cross-sectional survey. Nurse Educ. Pract. **69**, 103597 (2023). https://doi.org/10.1016/j.nepr.2023.103597
11. Li, H., Lu, F., Hou, M., Cui, K., Darbandi, M.: Customer satisfaction with bank services: the role of cloud services, security, e-learning and service quality. Technol. Soc. **64**, 101487 (2021). https://doi.org/10.1016/j.techsoc.2020.101487
12. Obidova, Z.: Methodological issues of physical education in secondary school. Inform. Econ. Manag. **2**(1), 0124–0131 (2023). https://doi.org/10.47813/2782-5280-2023-2-1-0124-0131

13. Sattarkulov, K.R.: Methods of studying the law of absolute radiation of a black body in physics courses of academic lyceums. Inform. Econ. Manag. **2**(1), 0132–0137 (2023). https://doi.org/10.47813/2782-5280-2023-2-1-0132-0137
14. Solomon, J., Wayne, N., Cowell, L., Stenson, S., Hubbard, G.P.: The development and use of e-learning modules to support care home staff caring for enterally tube fed patients. Clin. Nutr. ESPEN **48**, 520 (2022). https://doi.org/10.1016/j.clnesp.2022.02.103
15. Alanis, V.M., Recker, W., Ospina, P.A., Heuwieser, W., Virkler, P.D.: Dairy farm worker milking equipment training with an E-learning system. JDS Commun. **3**(5), 322–327 (2022). https://doi.org/10.3168/jdsc.2022-0217
16. Singh, S., Hussain, S.Z.: Mechanising E-learning for equiping start-up entrepreneurs. Mater. Today Proc. **37**(2), 2467–2469 (2021). https://doi.org/10.1016/j.matpr.2020.08.289
17. Malik, S., Rana, A.: E-Learning: role, advantages, and disadvantages of its implementation in higher education. Int. J. Inf. Commun. Comput. Technol. **8**(1), 403–408 (2020)
18. Tsarev, R., et al.: Gamification of the graph theory course. finding the shortest path by a greedy algorithm. In: Silhavy, R., Silhavy, P. (eds.) Networks and Systems in Cybernetics. CSOC 2023. Lecture Notes in Networks and Systems, vol. 723, pp. 209–216. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-35317-8_18
19. Lunev, D., Poletykin, S., Kudryavtsev, D.O.: Brain-computer interfaces: technology overview and modern solutions. Modern Innov. Syst. Technol. **2**(3), 0117–0126 (2022). https://doi.org/10.47813/2782-2818-2022-2-3-0117-0126
20. Fan, K., Liu, W., He, K., Wang, Z., Ou, S., Wu, Y.: Review: the application of artificial intelligence in distribution network engineering field. Inform. Econ. Manag. **2**(1), 0210–0218 (2023). https://doi.org/10.47813/2782-5280-2023-2-1-0210-0218
21. Ezaldeen, H., Bisoy, S.K., Misra, R., Alatrash, R.: Semantics aware intelligent framework for content-based e-learning recommendation. Nat. Lang. Process. J. **2023**, 100008 (2023). https://doi.org/10.1016/j.nlp.2023.100008
22. Deetjen-Ruiz, R., et al.: Applying ant colony optimisation when choosing an individual learning trajectory. In: Silhavy, R., Silhavy, P. (eds.) Networks and Systems in Cybernetics. CSOC 2023. Lecture Notes in Networks and Systems, vol. 723, pp. 587–594. Springer, Cham (2023).https://doi.org/10.1007/978-3-031-35317-8_53
23. Vedavathi, N., Kumar, A.K.M.: E-learning course recommendation based on sentiment analysis using hybrid Elman similarity. Knowl.-Based Syst. **259**, 110086 (2023). https://doi.org/10.1016/j.knosys.2022.110086
24. Gruzenkin, D.V., et al.: Neural networks to solve modern artificial intelligence tasks. J. Phys. Conf. Ser. **1399**(3), 033058 (2019). https://doi.org/10.1088/1742-6596/1399/3/033058
25. Samojlov, A.S., Goloborodko, E.V., Klyuchnikov, M.S.: Big data, machine learning and precision forecasting in sport medicine. Healthcare Educ. Secur. **1**(17), 7–17 (2019)
26. Semenenko, M.G., et al.: How to use neural network and web technologies in modeling complex technical systemsю. In: IOP Conference Series: Materials Science and Engineering, vol. 537, no. 3, p. 032095 (2019). https://doi.org/10.1088/1757-899X/537/3/032095
27. Semenova, E.A., Tsepkova, S.M.: Neural networks as a financial instrument. Inform. Econ. Manag. **1**(2), 0168–0175 (2022). https://doi.org/10.47813/2782-5280-2022-1-2-0168-0175
28. Elman, J.L.: Finding structure in time. Cogn. Sci. **14**(2), 179–211 (1990)
29. Li, L., Xie, X., Gao, T., Wang, J.: A modified conjugate gradient-based Elman neural network. Cogn. Syst. Res. **68**, 62–72 (2021). https://doi.org/10.1016/j.cogsys.2021.02.001
30. Toha, S.F., Tokhi, M.O.: MLP and Elman recurrent neural network modelling for the TRMS. In: Proceedings of the 7th IEEE International Conference on Cybernetic Intelligent Systems, pp. 1–6. IEEE, London, UK (2008). https://doi.org/10.1109/UKRICIS.2008.4798969
31. Zhang, Y., Wang, X., Tang, H.: An improved Elman neural network with piecewise weighted gradient for time series prediction. Neurocomputing **359**, 199–208 (2019). https://doi.org/10.1016/j.neucom.2019.06.001

32. Fan, Y., Yang, W: A backpropagation learning algorithm with graph regularization for feed-forward neural networks. Inf. Sci. **607**, 263−277 (2022)https://doi.org/10.1016/j.ins.2022.05.121
33. Glushchenko, A., Petrov, V., Lastochkin, K.: Backpropagation method modification using Taylor series to improve accuracy of offline neural network training. Procedia Comput. Sci. **186**, 202–209 (2021). https://doi.org/10.1016/j.procs.2021.04.139
34. Kaveh, A., Servati, H.: Design of double layer grids using backpropagation neural networks. Comput. Struct. **79**(17), 1561–1568 (2001). https://doi.org/10.1016/S0045-7949(01)00034-7
35. Mandischer, M.: A comparison of evolution strategies and backpropagation for neural network training. Neurocomputing **42**(1–4), 87–117 (2002). https://doi.org/10.1016/S0925-2312(01)00596-3
36. Zaras, A., Passalis, N., Tefas, A.: Deep Learning for Robot Perception and Cognition, Chapter 2 - Neural networks and backpropagation. Eds Iosifidis, A., Tefas, A. Academic Press (2022). https://doi.org/10.1016/B978-0-32-385787-1.00007-5

Empowering Islamic-Based Digital Competence and Skills: How to Drive It into Reconstructing Safety Strategy from Gender Violence

Miftachul Huda[1(✉)], Mukhamad Hadi Musolin[2], Anassuzastri Ahmad[2], Andi Muhammad Yauri[3], Abu Bakar[3], Muhammad Zuhri[3], Mujahidin[3], and Uswatun Hasanah[3]

[1] Sultan Idris Education University, Tanjung Malim, Malaysia
halimelhuda@gmail.com
[2] Sultan Abdul Halim Mu'adzam Shah International Islamic University, Tanjung Malim, Malaysia
[3] State Institute for Islamic Studies, Bone, South Sulawesi, Indonesia

Abstract. In the last decade, the urgent demand in empowering the digital competence and skills has been widely emerged amidst the digital society circumstance. This aim is to ensure its critical insights could be optimized to recon-structure safety strategy from gender violence. The digital competence and skills integration could be made through having a sufficient comprehension followed with practical stability commitment. The objective of this paper is to examine the how to do in empowering the strategic balance between digital competence and skills as an attempt to drive the pathway to lead to enhance safety strategy from gender violence. The systematic literature review was applied from the number of reviewed articles, including journals, books, proceeding and chapters related to the field. The finding reveals that empowering digital competence and skills to enhance reconstructing safety strategy from gender violence should go through expanding the knowledge comprehension, reflected practical skills and continued improvement commitment. This study is supposed to contribute in giving a strategic value on expanding the knowledge improvement in the digital competence and skills to improve safety strategy from gender violence. Moreover, this paper could be also recommended among the researchers, stakeholders and practitioners in the field of cyber security.

Keywords: digital competence · digital skills · knowledge comprehension · reflected practical skills · continued improvement and assessment and gender violence

1 Introduction

In the recent years, the massive expansion of digital-based human interaction across the various fields and sectors has been widely considered as the real impact of digital revolution industry. The full range of digital technologies could bring the information and

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 146–164, 2024.
https://doi.org/10.1007/978-3-031-53552-9_14

communication strategy with being a full confident sense on a critical use of an ultimate point of solving the digital issues in all life aspects [1, 59]. A significant contribution of considering a digital-based competence could be made through helping enhance the digital skills in underlying to master the wide range of related basic competences for instance language or communication skills [60, 61]. As this is a part of main gate to transmit the information relevant to the need and demand, the strategic basic skill requires to have a sufficient understanding of looking at the whole context of nature of forming the better competence [2, 62].

In line with the cyber-security sector, the element of what digital atmosphere could offer in performing the digital safety strategy from gender violence is being an important value in helping the real impact of digital revolution age as at this moment goes in [3, 63]. As a result, the development of digital competence has to be expanded through addressing information initiative of transmitting the data literacy in enabling the digital communities to communicate with an online collaboration [4, 64]. On this view, the continued commitment should be enlarged in creating the digital content with a safety in giving a strategic solving on the problem [65, 66]. With this regard, it is clear that the necessity of stressing the digital competence and skills in empowering the digital communication and skills would give a beneficial value in enhancing the digital community strength.

In addition, the strategic essence of having both sufficient comprehension and practical aspects played a significant role in advancing the digital community empowerment. Apart from the problematic issues raised such as insufficient knowledge of digital competence, lack of time management, instability of human life arrangement [5, 67], all these could lead to unproductivity for the digital users [68]. As such, the need to enhance the strategic approach of empowering digital competence and skills should be taken into consideration in the sense that played a pathway of driving into the reconstructing the digital safety strategy from gender violence [6]. At this point of view, this paper aims to examine the core point of distributing the digital competence and skills in going further on elaborating the confident responsibility in order to lead to achieve the digital atmosphere of safety strategy from gender violence.

2 Literature Review

2.1 Digital Competence and Skills

Digital competence refers to one of the key indicators of digital users in enabling them to have an interaction with their peers in a variety of purposes of life. The use of digital skills is potentially following the digital competence in grabbing the sufficient practice among the users [7, 69]. Through providing skilful abilities on adapting the sufficient digital, such this competence is being a key element to enhance practical implementation together with having a sufficient feeling on the confident engagement [70, 71]. With this regard, the more improvement on developing the use of technology at any purpose should be taken into consideration in advancing the information and communication technology (ICT) skills [8, 72]. In the purpose of teaching for instance, the adoption of bringing the technologies in underlying the teaching and learning process could be made in a proper manner. In further, the digital skills and competence in underlying

the teaching performance is required to have sufficient preparation together with the proper facilitation. As a result, providing this digital environment and space would lead to enhance the proper collaboration among their peers in underlying the learning.

In addition, the new innovation of strategic ways on skills and practices in the usage and adaptation of ICT to enhance the teaching might contribute to give a positive feedback on their task completion [73]. As a result, the essence of having a sufficient comprehension on the extent of technology skills aims to increase the digital teaching and learning practices [9, 74]. With this regard, getting an active participation in online skills should have a sufficient preparation on the digital platform arrangement [75, 76]. Moreover, the collaborative engagement created through adapting the materials with the peers is required to optimise the use of social networks on their educational purpose.

In line with advancing the wide elaboration on developing the digital competence, the early start could be made through deciding the appropriate selection of technology type in order to adapt and adopt in a proper manner [10, 77]). This is important that the type of tools for the education usage has to do with allocating the time which will be spent in assisting to complete their assignment. At this point of view, the allocating time and practice of digital tool enhancement should be carefully considered with a wise basis [11, 78]. In order to achieve this attainment, the strategic ways on adhering the principles in digital practice should be also committed in building the digital commitment. In terms of engaging the digital skills and practices on education purpose, the comprehension tool to gain a clear picture of the information has to come from the essence of communicating and interpreting what has been transmitted through such digital tool [12, 79]. It indicated that the continued development on digital skills and practices should bring along with an effective basis of technical format on technological tool adoption [80, 81]. Moreover, digital competence and skills are required to enhance the digital practice in underlying their education purpose at any level within the proper manner.

2.2 Digital Safety Strategy from Gender Violence

The essence of safety strategy from gender violence in digital age refers to encouraging strength of knowledge on how to do within achieving cyber-security lens in incorporating the social circumstance and networking space [13, 82]. The aim of the extent to transmitting both social networking and urban computing into designing and delivering the course arrangement is to expand the safety strategy from gender violence achievement. The consequent intersection in obtaining the essence of digital safety strategy from gender violence should come up with evolving the hyper-growth from adapting to adopting the digital technology within an online platform basis [14, 83]. As a result, the skill and practice on the way to search for and communicate with the students' peers played a significant role in obtaining the interactive circumstance. This is because the strategic essence of enhancing the safety strategy from gender violence is particularly a continuing process with a hands-on way in transmitting the knowledge delivery into the peers and members [84, 85]. Concerning transfer the knowledge, getting an active participation through listening and responding the instruction process is being a particular phase to do with obtaining the assigned material [15, 86]. As such, the condition to achieve the safety strategy from gender violence becomes an intervention point in gaining an active engagement among the digital users.

In addition, the safety strategy from gender violence in digital age is pointed out a comprehensive pathway comprising the multi-dimension of strategic approach and materials. The committed awareness of expanding the strategic approach on digital safety strategy from gender violence is positively enhanced with combining the multiple sides of medium of instruction such as audio and visual aspect [16, 87]. As a result, the continuing process of making the animation using video, images, text and audio could be determined through bringing the critical inquiry inspection to lead to the safety environment. In order to gaining a usual aim for self-regulated inquiry process, the delivery on transmitting an online sources through the wide range of platforms on networking basis [17, 88]. The support of enhancing the safety strategy from gender violence could be made through providing the instructional process to supplement an active engagement as the central point of integrating safety strategy from gender violence enhancement. The number of activity on such support service is required to enhance the active engagement in arranging the time planning and management [18, 89]. It is important to expand the necessity of taking a beneficial value on actualising the digital safety strategy from gender violence through providing the interactive resources for the cyber-security purpose.

In line with allowing an active engagement to enable the digital safety strategy from gender violence, the strategic arrangement on planning the extent of time management should do with having a sufficient capacity of the digital space with the scalable effective and efficient scale. On this view, the flexibility achievement on the digital teaching competence could be made through the session of training program among the educators [19, 90]. It is important to take note that ensuring the learning inquiry process with the digital interactive circumstance is pointed out having a baseline of knowledge comprehension about the particular content. The initial understanding is that the wide range of levels on the prior knowledge should be incorporated through gaining the factual aspects of what has been taught with a straighter forward [20, 91]. This is to ensure the clarity of subject and content delivered into achieving the safety strategy from gender violence. As such, the preparation could be made through having the availability to contact time to respond the contested or complex ideas in enabling and facilitating the work collaboration and consolidation [21, 92]. The strategic essence on revising the digital interactive aspects has to bring along with the social life instruction design together with addressing the time table arrangement in supporting a safety strategy from gender violence.

3 Methodology

The process of highlighting the steps taken as flow of applying for the systematic literature review of the latest literature on digital competence and skills in driving into reconstructing digital safety strategy from gender violence. The assessment was made to assist the review process in revealing the positive feedback of digital competence and skills in order to underlie the cyber-security achievement with a safety circumstance. With the wide range of achieving the safety environment outcomes within the cyber-security arrangement, the strategic review process might help in highlighting the initial point as the potentials to improve the safety efficiency amidst the digital medium. The arrangement to capitalize the mobile devices is pointed out providing the safety achievement with an interactive essence of digital media aspect. In giving insights into providing

the detail description, the systematic literature review is applied with researching to focus on selecting the appropriate and proper data as the essential phase.

The objective of this paper refers to examine the how to do in empowering the strategic balance between digital competence and skills as an attempt to drive the pathway to lead to enhance digital safety strategy from gender violence. In order to achieve this, both skills and experiences are regulated with the principles in digital activities, in encountering the main point of emerging trends on social media and technology. As such, systematic literature review was critically conducted from peer reviewed articles from journals, proceedings and books. By considering the way to adopt, implement and then utilize such media technology, the proper manner is supposed to give insights in transmitting the wise approach on information data. Moreover, the cross-review was made through addressing the diverse backgrounds of users in terms of culture and life ways. The critical analysis was also made by conducting the keywords such as digital information, trust and transparency and information quality amidst the new norm of digital circumstance. As a result, creating the safe space on digital space is mainly the initial aim to enhance the transparency and trust amidst the society culture.

4 Analysis and Discussion

4.1 Driving Pathway of Digital Competence and Skills for Safety Strategy from Gender Violence

Digital safety strategy from gender violence refers to the skills in practising the instructional process with the technological platform support. The continuous development of expanding both digital competence and skills is concisely elaborated to underlie the strategic process to lead to the safety engagement [22, 93]. With this regard, the strategic essence of applying for the digital competence and skills should bring along with driving the pathway as the main achievement plan. The competence on driving the digital interactive essence for the purpose of safety inquiry process has to be started with the digital skills and practice enhancement and framework [23, 94]. The potential inclusion of determining the strategic pathway on capturing the digital environment space is pointed out expanding the competence and confidence strength. This aims to ensure the early phase until the completion process could be within the proper pathway [24, 95]. In strategizing the clear picture of delivering online sources, the main achievement of this task refers to do with describing the graphic essence apart from the textual illustration.

In addition, the strategic way on making confident in enabling the safety strategy from gender violence for being more interactive circumstance should be taken into consideration in a proper manner collaborated between technological competence and pedagogical experts [96, 97]. The strategic expansion of bringing the future learning enhancement should do with having the well-arranged framework on digital interactive essence [25]. With this regard, educating the way to do for safety concern enhancement with digital support is concisely engaged to transmit into the preliminary assessment in developing the cyber-security environment. Providing the advancement to empower the digital materials and online platform is required to support the attempts in managing the digital tool use [26]. As a result, the continued improvement through a proper assessment of digital safety strategy from gender violence should be empowered in facilitating

the essence of digital competence [98, 99]. Constructing this initiative in enhancing the educational context has to specifically follow the strategic pathway on performing quality in underlying the safety strategy from gender violence process amidst the digital platform.

In line with expanding the wide range of applying for the safety strategy from gender violence with technology, the strategic basis in conceptualizing the content into skills and experts would lead to give a critical insight into building the safety concern amidst online atmosphere [100, 101]. From this point of view, the safety strategy from gender violence with an interactive circumstance could be carried out through such as augmented reality, smartphone engagement, safety strategy from gender violence with Social Networking Site (SNS) environment [27, 28]. In order to come up with this attainment, the related instrument on adapting and adopting the digital platform-instructional process refers to give an instrumental analysis involved in the primary classroom arrangement [102, 103]. Based on this reality, the setting of creating the digital safety strategy from gender violence in the capacity to manage the process from the beginning to the end would give a beneficial value in transmitting the strategic concern on safety atmosphere and instruction process [28, 104]. On this view, the further elaboration on expanding the digital instruction involved in the usage of software development program has to be taken into consideration in a proper manner. It is important to have a clear picture of mapping the adaptation followed by adopting the digital platform to enhance a safety strategy from gender violence environment.

4.2 Adapting Digital Competence and Practical Skills for Safety Strategy from Gender Violence

The further commitment to adopt the digital competence should be initiated with having a clear inspection to adapt the competence and practical skills. As a result, the critical exploration to gain an involvement of knowledge comprehension on what to do in line with the use of digital platform in higher education context for instance should be incorporated in carrying out the cyber-security improvement [48, 105]. The advancement in making an impact to the safety strategy orientation could be concisely pointed out improving the digital environment-enabled strategic inquiry process [29, 106]. The online environments with an online aspects in providing the raising value of interactivity would lead to give a supportive strength for the safety strategy from gender violence space among the digital users. With this regard, the engagement in enabling the digital users to have a sufficient sight to adapt the visualization on both content and object is widely coordinated with an active interaction in the response towards the digital instruction process [30, 35]. Through getting ready with optimizing the dimension from visual, audio and mixing value, the continued exploration in adapting such environment condition has to do with navigating a strategic pathway in enhancing and interacting the digital users to get involved in the online-based social interaction.

In addition, the digital competence would be significant when the simultaneous capacity of both cognitive and affective skills to give an insightful value in providing a clear picture of planning management and achievement. As a result, the critical insight to understand the digital atmosphere should follow the features and characteristics in giving an outstanding value to the safety strategy from gender violence process [31]. The

outstanding feedback which can be gained is to contribute in strengthening instructional process of digital online interaction process [107, 108]. The wide range of digital competence and skills could be integrated to achieve the safety strategy from gender violence in the context of cyber-security management pathway. With this regard, the strategic effort to educate by the usage of online instruction would show the initiative in supporting the cyber-security enhancement [31]. On this view, the digital online technology-enabled initiative in contributing the workflow on upgrading the conventional and traditional-based social instruction would give a significant impact in educating the pathway of safety transformation. At this point of view, the significant way in transforming the learning pathway into digital interactive basis requires to posit in enabling the main aim of efficient and effective achievement [109, 110]. The primary focus of this impactful outcome would lead to commit with the digital technology adaptation and adoption so that the strategic attempt in forming the interactive atmosphere.

In line with obtaining the strategic effort with the attempts to make traditional value of instructional processes to become more efficient, adapting the proper way of paradigm for safety strategy from gender violence is being a transformative application to integrate with the digital environment landscape. As a result, the practical identification on suggesting the transformative applications in empowering the safety strategy from gender violence with digital technology is applied for the partnership engagement [32, 111]. The essence of adapting the transformation pathway in applying for the digital interaction capacity would become the main orientation of being partner engagement in underlying the online-based social interaction. In empowering the extent of safety achievement process, the combination of physical and virtual based instruction is required to provide the technical skills exploration [33, 112]. The model in helping the safety enhancement in a meaningful way could be empowered in continuing the advancement of cyber-security inquiry process. With a sufficient comprehension on the security capacity, both theoretical concept and practical capability should be accommodated in looking into detail about the way of digital safety strategy from gender violence enhancement [34]. As such, maintaining the realization of both individual capacity and professional experts in underlying digital safety strategy from gender violence is significant to sustain the necessary act of continuing adaptation on critical reflection development [113, 114]. The main orientation of gaining the existing views of digital interactive enhancement on online social interaction process is important to enable the cyber-security process on the subject content.

4.3 Empowering Digital Competence and Skills for Safety Strategy from Gender Violence

The essence of empowering digital competence and skills refers to the strategic attempts to develop the technology basis with online virtual instrument in underlying instructional design of cyber-security management. The advancement of digital space on obtaining safety strategy from gender violence is widely seen to upgrade traditional basis from face to face classroom to hybrid mixture [35, 115]. In order to achieve this attainment, the strategic approach of digital environment for safety strategy from gender violence space could be widely engaged into distributing the lens of transforming the cyber-security sector with the full technology support. As a result, the thorough approach to

underlie the transformation process in the safety concern and enhancement should do with empowering the digital competence as an attempt to advance the online environment potentials [36, 116]. At this point of view, the need to respond the potentials of upgrading the conventional approach in the context of online social interaction into the cyberspace-supported instructional medium has to be concisely taken into consideration. With this regard, the safety strategy from gender violence should start with having a sufficient knowledge comprehension articulated into the well-designed plan [37, 117].

In addition, the digital competence and skill is considered to foster the strategic essence of transformative safety concern within the online social interaction environment. In order to have a sufficient preparation in gaining the successful achievement, the strength should be made by recommending the strategy properly in line with building the conducive environment [38, 118]. The strategic attainment on creating a safe-based circumstance is concisely engaged to support the environment reflected to encourage the safety inquiry process with a sufficiently interactive basis. It is important to note that having the pathway to underlie the digital users' thinking reflected to support the experiences could be made through a sufficient balance between beliefs and commitment. With this regard, the strategic use of social interaction technology in promoting the active users engagement refers to expand the continuous development on instruction design pathway [39, 119]. In order to have such arrangement to be in reality circumstance, the strategic management on looking into detail about the challenging problems is required to initially address the individual capacity in assisting the practical oriented solutions [120]. On this view, the need to help apply for the cyber-security in action amidst the digital online circumstance should be taken into consideration in a proper manner of strategizing the attempts to utilize technology support transmitted in the education context.

In line with empowering digital competence and skills, the strategic support of how to advance the essence of digital safety strategy from gender violence with both effective and efficient outcome is actively being fostered in transforming the practice and experiences among the digital users [121, 122]. In order to give a portion of safety strategy from gender violence enhancement, the practical skills in both online and conventional classroom could be incorporated such as communication skills, listening skills through an audio platform [40]. The individual assessment and feedback encouragement are the central point on looking into detail about the digital safety strategy from gender violence enhancement. Moreover, the interactive instruction process and material resources could be also the encouraging aspects on supporting the digital safety strategy from gender violence environment [41]. Such this initiative might also be considered as the strategic essence of modern learning environment (MLE) with big data support model framework. As a result, the empowerment attempt to be made through having a clear picture and mapping of instructions by adapting and adopting the appropriate technology support [42, 123]. The initial pathway on developing the instructional process of safety strategy from gender violence should do with helping the strategic point of ideas reflected into the practice with an effective and efficient instruction design.

4.4 Expanding Knowledge Comprehension for Safety Strategy from Gender Violence

The attempts to expand the knowledge improvement in the digital competence and skills refer to search for the detailed information on how to do in line with advancing a safety strategy from gender violence supported by the digital environment. The way to do in contributing to give a strategic value in improving the digital safety strategy from gender violence could be incorporated in recommending the further elaboration amongst the direct involved-implementer, learners and educators [43, 124]. In order to give a proper format in developing their knowledge comprehension, digital users' skills and competence should be initiated in focusing on the online interaction potentials in the inquiry process. The main element in contributing the wider engagement to accept the digital environment is to do considerably with the instructional instrument in advancing effectiveness and efficiency in creating safety environments with being more enjoyable and immersive [44, 125].

In addition, it is necessary to give the cyber-security enhancement plan to be well arranged to give more efficaciously in the digital interaction context. As a result, digital interactive-based support by the real-world environments could be strengthened by embedding the mixture between physical and virtual contents [45, 126]. The main value is to give a full support on the inquiry process of helping improve digital interaction from advancing the conceptual understanding followed by the cyber-security inquiry with the support of technology tool. In further, the practical commitment on utilizing the digital interactive contents and also internet resources would enable digital users in performing their assignment with being more flexible and convenient [2, 46]. From this view, the necessary act to adapt both theoretical and practical balance in increasing the digital users' activity and motivation should be formed in regulating the digital multimedia supported instruction setting. Through adopting the use of mobile devices, material resources of digital safety strategy from gender violence could be developed with being more accessible to the digital users in the borderless space basis [47, 127]. Apart from conditioning the geographical and temporal barriers, the further detailed description on carrying out the social media-enabled environment has to focus on digital assisted social practices to construct the interactive essence.

Collaborated with the strategic pathway on building the systematic formulation on applying for the appropriate approach, the current assessment on determining the digital interactive-supported environment should come up searching for the up to date applications in the cyber-security context. As a result, the attempt to carry out highlighting both positive feedback and negative one in terms of real application has to be incorporated in implementing both adaptation and adoption of relevant multimedia elements [48, 128]. In terms of determining the formulation on advancing instructional design on digital safety strategy from gender violence environment, the strategic approach in applying for the achievement plan should be well designed with the instrumental tool capacity. In terms of continuing the individual capacity in developing the knowledge expansion on comprehending the digital safety strategy from gender violence element, the particular approach in assessing the proper application on this initiative is potentially being an ultimate transmission of looking into detail about the wide range of important purposes in safety orientation context [49]. The further detailed identification should do

with incorporating both beneficial and potential value in implementing the technology-mediated instructional design. As such, the further expansion from adapting and adopting the media literacy-supported learning inquiry process in giving a critical insight into serving as an encouraging element leading to effective and efficient achievement.

4.5 Strengthening Practical Skills as Continued Improvement for Safety Strategy from Gender Violence

The value of continuing the development of digital safety strategy from gender violence should come up with building the supportive instrument through practical skills on achieving such plan arrangement. Through advancing the supporting tool in enabling the digital interaction pathway, the main purpose aims to enhance the continued improvement of creating more efficient and effective circumstance of digital safety strategy from gender violence enhancement [50, 59]. Moreover, the attempts to strengthen practical skills could be performed through an active engagement in adopting the cyber-security inquiry process. As a result, the occupation to achieve digital interactive-supported instruction circumstance is aimed at building usefulness together with building the novel technology-enabled digital user potentials [51, 129]. In consequence, the more update of assessing the related sources is concisely carried in approaching the properly suitable method. This point is to give a continuing support in highlighting the feedback of the cyber-security performance with being more motivated articulated into perceived attitudes and particular attention.

In revealing such achievement, the need to incorporate the education with the interactive technology should be taken into consideration in an active implementation of an active setting to further enhance the safety achievement process through interactive and immersive environments [52, 130]. In further, achieving the beneficial value to achieve among digital users could be performed in advancing the proper applications of digital-supported social interaction management [53, 131]. Moreover, the multimedia and media literacy education in classroom settings process should be performed with early phase of determining the proper strategy in the basis of technology enhancement. With this regard, expanding the category on analysing the further elaboration of featured criteria in grouping the characteristics of digital interactive enabled practical skills is potentially leading to enhance the essential value of applying for the continued improvement [54, 132]. Through highlighting this initiative transmitted into the specific application, the performance of strengthening practical skills on digital safety strategy from gender violence enhancement should be taken into consideration in particular in line with a proper manner.

In addition, the further exploration on strengthening the practical skills as continued improvement in digital safety strategy from gender violence comes to be initially initiated to ensure the learning environment [133, 134]. The more knowledge comprehension on applying for the digital skills could be incorporated through a sufficient learning expansion with a useful applicability for interactive engagement [55, 135]. In enabling the digital safety strategy from gender violence environment, it is important to have a balance on visualizing the outcome based digital interaction in helping the improvement of thinking pathway skills. As a result, the systematic process on achieving the main point of practical skills should do with highlighting the detailed description of actively applying

for the digital environment space [56]. With this regard, the strategic applications of digital environment on underlying the cyber-security contexts could be continued in the improvement of fulfilling the featured characteristics of an online space. At this point of view, the main orientation aims to assist the safety inquiry achievement with the digital enabled instrumental design in order to gain an active engagement on the entire process [57]. As such, the way to enhance practical stability on digital competence and skills is being an ultimate point to lead to a thorough process in ensuring the digital users' better comprehension [58]. Thus, such this initiative is followed by a commitment to carry out the rules, procedures and plan arrangement.

5 Conclusion

In the last decade, the urgent demand in empowering the digital competence and skills has been widely emerged amidst the digital society circumstance. This aim is to ensure its critical insights could be optimised to recon-structure digital safety strategy from gender violence. The objective of this paper refers to examine the how to do in empowering the strategic balance between digital competence and skills as an attempt to drive the pathway to lead to enhance digital safety strategy from gender violence. The systematic literature review was applied from the number of reviewed articles, including journals, books, proceeding and chapters related to the field. The finding reveals that empowering digital competence and skills to enhance reconstructing digital safety strategy from gender violence should go through expanding the knowledge comprehension, reflected practical skills and continued improvement and assessment. This study is supposed to contribute in giving a strategic value on expanding the knowledge improvement in the digital competence and skills to improve digital safety strategy from gender violence. Moreover, this paper could be also recommended among the researchers, stakeholders and practitioners in the cyber-security.

References

1. Almås, A.G., Bueie, A.A., Aagaard, T.: From digital competence to professional digital competence: student teachers' experiences of and reflections on how teacher education prepares them for working life. Nordic J. Compar. Int. Educ. (NJCIE) **5**(4), 70–85 (2021)
2. Anshari, M., Almunawar, M.N., Shahrill, M., Wicaksono, D.K., Huda, M.: Smartphones usage in the classrooms: Learning aid or interference? Educ. Inf. Technol. **22**(6), 3063–3079 (2017)
3. Bartolomé, J., Garaizar, P., Larrucea, X.: A Pragmatic Approach for Evaluating and Accrediting Digital Competence of Digital Profiles: A Case Study of Entrepreneurs and Remote Workers. Technology, Knowledge and Learning, pp. 1–36 (2021)
4. Beardsley, M., Albó, L., Aragón, P., Hernández-Leo, D.: Emergency education effects on teacher abilities and motivation to use digital technologies. Br. J. Edu. Technol. **52**(4), 1455–1477 (2021)
5. Behnamnia, N., Kamsin, A., Ismail, M.A.B., Hayati, A.: The effective components of creativity in digital game-based learning among young children: a case study. Child Youth Serv. Rev. **116**, 105227 (2020)

6. Cabezas-González, M., Casillas-Martín, S., García-Peñalvo, F.J.: The digital competence of pre-service educators: the influence of personal variables. Sustainability **13**(4), 2318 (2021)
7. Cattaneo, A.A., Antonietti, C., Rauseo, M.: How digitalised are vocational teachers? Assessing digital competence in vocational education and looking at its underlying factors. Comput. Educ. **176**, 104358 (2022)
8. Colás-Bravo-Bravo, P., Conde-Jiménez, J., Reyes-de-Cózar, S.: The development of the digital teaching competence from a sociocultural approach. Comunicar Media Educ. Res. J. **27**(61), 21–32 (2019). https://doi.org/10.3916/C61-2019-02
9. Colás-Bravo, P., Conde-Jiménez, J., Reyes-de-Cózar, S.: Sustainability and digital teaching competence in higher education. Sustainability **13**(22), 12354 (2021)
10. Darazha, I., Lyazzat, R., Ulzharkyn, A., Saira, Z., Manat, Z.: Digital Competence of a Teacher in a Pandemic. In: 2021 9th International Conference on Information and Education Technology (ICIET), pp. 324–328. IEEE, March 2021
11. Galindo-Domínguez, H., Bezanilla, M.-J.: Promoting time management and self-efficacy through digital competence in university students: a mediational model. Contemp. Educ. Technol. **13**(2), ep294 (2021). https://doi.org/10.30935/cedtech/9607
12. Galindo-Domínguez, H., Bezanilla, M.J.: Digital competence in the training of pre-service teachers: Perceptions of students in the degrees of early childhood education and primary education. J. Safety Strategy Gender Violence Teacher Educ. **37**(4), 262–278 (2021)
13. Gordillo, A., Barra, E., Garaizar, P., López-Pernas, S.: Use of a simulated social network as an educational tool to enhance teacher digital competence. IEEE Revista Iberoamericana de Tecnologias del Aprendizaje **16**(1), 107–114 (2021)
14. Guillén-Gámez, F.D., Mayorga-Fernández, M., Bravo-Agapito, J., Escribano-Ortiz, D.: Analysis of teachers' pedagogical digital competence: identification of factors predicting their acquisition. Technol. Knowl. Learn. **26**(3), 481–498 (2021)
15. Guillén-Gámez, F., Cabero-Almenara, J., Llorente-Cejudo, C., Palacios-Rodríguez, A.: Differential analysis of the years of experience of higher education teachers, their digital competence and use of digital resources: comparative research methods. Technology, Knowledge and Learning, pp. 1–21 (2021)
16. Hanafi, H.F., Wahab, M.H.A., Selamat, A.Z., Masnan, A.H., Huda, M.: A systematic review of augmented reality in multimedia learning outcomes in education. In: Singh, M., Kang, D.-K., Lee, J.-H., Tiwary, U.S., Singh, D., Chung, W.-Y. (eds.) Intelligent Human Computer Interaction: 12th International Conference, IHCI 2020, Daegu, South Korea, November 24–26, 2020, Proceedings, Part II, pp. 63–72. Springer International Publishing, Cham (2021). https://doi.org/10.1007/978-3-030-68452-5_7
17. Huda, M., Hashim, A.: Towards professional and ethical balance: insights into application strategy on media literacy education. Kybernetes **51**(3), 1280–1300 (2021). https://doi.org/10.1108/K-07-2017-0252
18. Huda, M., Rosman, A.S., Mohamed, A.K., Marni, N.: Empowering adaptive learning technology: practical insights from social network site (SNS). In: Silhavy, R., Silhavy, P., Prokopova, Z. (eds.) Software Engineering Application in Informatics: Proceedings of 5th Computational Methods in Systems and Software 2021, Vol. 1, pp. 410–420. Springer International Publishing, Cham (2021). https://doi.org/10.1007/978-3-030-90318-3_34
19. Huda, M., Teh, K.S.M.: Empowering professional and ethical competence on reflective teaching practice in digital era. In: Dikilitas, K., Mede, E., Atay, D. (eds.) Mentorship Strategies in Teacher Education, pp. 136–152. IGI Global (2018). https://doi.org/10.4018/978-1-5225-4050-2.ch007
20. Iglesias-Rodríguez, A., Hernández-Martín, A., Martín-González, Y., Herráez-Corredera, P.: Design, validation and implementation of a questionnaire to assess teenagers' digital competence in the area of communication in digital environments. Sustainability **13**(12), 6733 (2021)

21. Isoda, M., Estrella, S., Zakaryan, D., Baldin, Y., Olfos, R., Araya, R.: Digital competence of a teacher involved in the implementation of a cross-border lesson for classrooms in Brazil and Chile. Int. J. Lesson Learn. Stud. **10**, 362–377 (2021)
22. Kalimullina, O., Tarman, B., Stepanova, I.: Education in the context of digitalization and culture: evolution of the teacher's role, pre-pandemic overview. J. Ethnic Cult. Stud. **8**(1), 226–238 (2021)
23. Khalili, H.M., Rosman, A.S., Mohamed, A.K., Marni, N.: Safety strategy from gender violence Enhancement Through Social Network Site (SNS). In: Silhavy R., Silhavy P., Prokopova Z. (eds) Software Engineering Application in Informatics. CoMeSySo 2021. Lecture Notes in Networks and Systems, vol 232. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-90318-3_35
24. Kilic, F., Karakuş, I.: New features of learners in education: digital awareness, digital competence, and digital fluency. In: Improving Scientific Communication for Lifelong Learners, pp. 113–132. IGI Global (2021)
25. Lohr, A., et al.: On powerpointers, clickerers, and digital pros: Investigating the initiation of safety strategy from gender violence activities by teachers in higher education. Comput. Hum. Behav. **119**, 106715 (2021)
26. Mehrvarz, M., Heidari, E., Farrokhnia, M., Noroozi, O.: The mediating role of digital informal learning in the relationship between students' digital competence and their academic performance. Comput. Educ. **167**, 104184 (2021)
27. Mishra, D., Sain, M.: Role of digital education to curb gender violence. In: Strategies for e-Service, e-Governance, and Cybersecurity: Challenges and Solutions for Efficiency and Sustainability, vol. 33 (2021)
28. Nasr, H.A.: Competences in digital online media literacy: towards convergence with emergency remote EFL learning. Int. J. Media Inf. Liter. **5**(2), 164–175 (2020)
29. Olesika, A., Lama, G., Rubene, Z.: Conceptualization of digital competence: perspectives from higher education. Int. J. Smart Educ. Urban Soc. (IJSEUS) **12**(2), 46–59 (2021)
30. Örtegren, A.: Digital citizenship and professional digital competence — Swedish subject teacher education in a PostDigital era. Postdigit. Sci. Educ. **4**(2), 467–493 (2022). https://doi.org/10.1007/s42438-022-00291-7
31. Passey, D., Bottino, R., Lewin, C., Sanchez, E. (eds.): Empowering Learners for Life in the Digital Age. IAICT, vol. 524. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-23513-0
32. Pérez-Calderón, E., Prieto-Ballester, J.M., Miguel-Barrado, V.: Analysis of digital competence for Spanish teachers at pre-university educational key stages during COVID-19. Int. J. Environ. Res. Public Health **18**(15), 8093 (2021)
33. Pérez-Navío, E., Ocaña-Moral, M.T., Martínez-Serrano, M.D.C.: University graduate students and digital competence: are future secondary school teachers digitally competent? Sustainability **13**(15), 8519 (2021)
34. Petrushenko, Y., Onopriienko, K., Onopriienko, I., Onopriienko, V.: Safety strategy from gender violence for adults in the context of education market development. In: 2021 11th International Conference on Advanced Computer Information Technologies (ACIT), pp. 465–468. IEEE, September 2021
35. Zainul, Z., Rasyid, A., Nasrul, W.: Distance learning model for islamic religious education subjects in non internet server provider (ISP) areas. Firdaus J. **1**(1), 45–53 (2021)
36. Rahman, M.F.F.A., Shafie, S., Mat, N.A.A.: Penerapan Konsep QEI Dan Teknologi dalam Pembangunan Kit Letti bagi Topik Integer Matematik Tingkatan 1. Firdaus J. **2**(2), 13–25 (2022)
37. Quaicoe, J.S., Pata, K.: Teachers' digital literacy and digital activity as digital divide components among basic schools in Ghana. Educ. Inf. Technol. **25**(5), 4077–4095 (2020)

38. Rahayu, N.W., Haningsih, S.: Digital parenting competence of mother as informal educator is not inline with internet access. Int. J. Child-Comput. Interact. **29**, 100291 (2021)
39. Embong, A.H., Yasin, M.F.M., Ghazaly, M.: Keberkesanan Modul Ulul Albab secara atas talian: Satu kajian di Universiti Malaysia Terengganu. Firdaus J. **1**(1), 93–102 (2021)
40. Selamat, A.Z., Adnan, M.A., Hanafi, H.F., Shukor, K.A.: Pengamalan Nilai Takwa sebagai Elemen Pembinaan Jati Diri dalam Kalangan Mahasiswa Muslim Di IPT. Firdaus J. **1**(1), 103–112 (2021)
41. Reisoğlu, İ.: How Does Digital Competence Training Affect Teachers' Professional Development and Activities? Technology, Knowledge and Learning, pp. 1–28 (2021)
42. Robles Moral, F.J., Fernández Díaz, M.: Future primary school teachers' digital competence in teaching science through the use of social media. Sustainability **13**(5), 2816 (2021)
43. Sailer, M., Murböck, J., Fischer, F.: Safety strategy from gender violence in schools: what does it take beyond digital technology? Teach. Teach. Educ. **103**, 103346 (2021)
44. Sailer, M., et al.: Technology-related teaching skills and attitudes: validation of a scenario-based self-assessment instrument for teachers. Comput. Hum. Behav. **115**, 106625 (2021)
45. Huda, M., et al.: Learning quality innovation through integration of pedagogical skill and adaptive technology. Int. J. Innov. Technol. Explor. Eng. **8**(9S3), 1538–1541 (2019)
46. Yuan, C., Wang, L., Eagle, J.: Empowering English language learners through digital literacies: research, complexities, and implications. Media Commun. **7**(2), 128–136 (2019)
47. Nurdani, N., Ritonga, M., Mursal, M.: Mastery learning as learning model to meet the passing grade of Al-Qur'an Hadith Subject at Madrasah Ibtidaiyah Negeri 4 Padang Pariaman. Firdaus J. **1**(1), 1–11 (2021)
48. Syofiarti, S., Riki, S., Ahmad, L., Rahmi, R.: The use of audiovisual media in learning and its impact on learning outcomes of Islamic cultural history at Madrasah Tsanawiyah Negeri 4 Pasaman. Firdaus J. **1**(1), 36–44 (2021). https://doi.org/10.37134/firdaus.vol1.1.4.2021
49. Mahruf, M., et al.: Emergency remote teaching and learning: Digital competencies and pedagogical transformation in resource-constrained contexts. In: Rezaul Islam, M., Behera, S.K., Naibaho, L. (eds.) Handbook of Research on Asian Perspectives of the Educational Impact of COVID-19:, pp. 175–200. IGI Global (2022). https://doi.org/10.4018/978-1-7998-8402-6.ch011
50. Shonfeld, M., et al.: Learning in digital environments: a model for cross-cultural alignment. Educ. Tech. Res. Dev. **69**(4), 2151–2170 (2021)
51. Sillat, L.H., Tammets, K., Laanpere, M.: Digital competence assessment methods in higher education: a systematic literature review. Educ. Sci. **11**(8), 402 (2021)
52. Spante, M., Hashemi, S.S., Lundin, M., Algers, A.: Digital competence and digital literacy in higher education research: systematic review of concept use. Cogent Educ. **5**(1), 1519143 (2018)
53. Willermark, S.M.J., Gellerstedt, M.: Digitalization, distance education, virtual classroom, high school, digital competence, COVID-19, ideal-type analysis. J. Educ. Comput. Res. 07356331211069424 (2022)
54. Wong, K.M., Moorhouse, B.L.: Digital competence and online language teaching: Hong Kong language teacher practices in primary and secondary classrooms. System **103**, 102653 (2021)
55. Huda, M., et al.: From digital ethics to digital partnership skills. In: Almunawar, M.N., Ordóñez, P., de Pablos, M., Anshari, (eds.) Digital Transformation for Business and Society: Contemporary Issues and Applications in Asia, pp. 292–310. Routledge, London (2023). https://doi.org/10.4324/9781003441298-14
56. Huda, M., et al.: Towards digital servant leadership for organisational stability: driving processes in the pandemic age. In: Digital Transformation for Business and Society, pp. 23–42. Routledge (2024)

57. Huda, M.: Trust as a key element for quality communication and information management: insights into developing safe cyber-organisational sustainability. Int. Jo. Organiz. Anal. (2023)
58. Huda, M.: Between accessibility and adaptability of digital platform: investigating learners' perspectives on digital learning infrastructure. High. Educ. Ski. Work. Based (2023)
59. Huda, M., Shahrill, M., Maseleno, A., Jasmi, K.A., Mustari, I., Basiron, B.: Exploring adaptive teaching competencies in big data era. Int. J. Emerg. Technol. Learn. **12**(3), 68–83 (2017)
60. Mohamad Shokri, S.S., Salihan, S.: Modul pembangunan Program Huffaz Profesional Universiti Tenaga Nasional: Satu pemerhatian kepada konstruk pembangunan Al-Quran. Firdaus J. **3**(2), 12–23 (2023). https://doi.org/10.37134/firdaus.vol3.2.2.2023
61. Berlian, Z., Huda, M.: Reflecting culturally responsive and communicative teaching (CRCT) through partnership commitment. Educ. Sci. **12**(5), 295 (2022)
62. Huda, M.: Towards an adaptive ethics on social networking sites (SNS): a critical reflection. J. Inf. Commun. Ethics Soc. **20**(2), 273–290 (2022)
63. Omar, M.N.: Inovasi pengajaran & pemudahcaraan menggunakan Aplikasi Ezi-Maq (MAHARAT AL-QURAN) untuk menarik minat pelajar menguasai Ilmu Tajwid. Firdaus J. **2**(2), 79–87 (2022). https://doi.org/10.37134/firdaus.vol2.2.8.2022
64. Huda, M., Sutopo, L., Liberty, Febrianto, Mustafa, M.C.: Digital information transparency for cyber security: critical points in social media trends. In: Arai, K. (eds.) Advances in Information and Communication. FICC 2022. LNNS, vol. 439, pp. 814–831. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-98015-3_55
65. Borham, A.H., et al.: Information and communication ethics in social media for indigenous people's religious understanding: a critical review. In: Proceedings of World Conference on Information Systems for Business Management: ISBM 2023. Springer, Cham (2024)
66. Ali, A.H.: Pelajar Khalifah Profesional Tempaan Ulul Albab sorotan penerapannya berasaskan komponen QEI di UPSI. Firdaus J. **3**(2), 51–63 (2023). https://doi.org/10.37134/firdaus.vol3.2.5.2023
67. Jusoh, A., Huda, M., Abdullah, R., Lee, N.: Development of digital heritage for archaeovisit tourism resilience: evidences from E-Lenggong web portal. In: Proceedings of World Conference on Information Systems for Business Management: ISBM 2023. Springer, Cham (2024)
68. Rahim, M.M.A., Huda, M., Borham, A.H., Kasim, A.Y., Othman, M.S.: Managing information leadership for learning performance: an empirical study among public school educators. In: Proceedings of World Conference on Information Systems for Business Management: ISBM 2023. Springer, Cham (2024)
69. Hehsan, A., et al.: Digital Muhadathah: framework model development for digital Arabic language learning. In: Cyber Security and Applications: Proceedings of ICTCS 2023, vol. 3. Springer, Cham (2024)
70. Muharom, F., Farhan, M., Athoillah, S., Rozihan, R., Muflihin, A., Huda, M.: Digital technology skills for professional development: insights into quality instruction performance. In: ICT: Applications and Social Interfaces: Proceedings of ICTCS 2023, vol. 1. Springer, Cham (2024)
71. Muhamad, N., Huda, M., Hashim, A., Tabrani, Z.A., Maarif, M.A.: Managing technology integration for teaching strategy: public school educators' beliefs and practices. In: ICT: Applications and Social Interfaces: Proceedings of ICTCS 2023, vol. 1. Springer, Cham (2024)
72. Wahid, A., et al.: Digital technology for indigenous people's knowledge acquisition process: insights from empirical literature analysis. In: Intelligent Strategies for ICT: Proceedings of ICTCS 2023, vol. 2. Springer, Cham (2024)

73. Huda, M., Taisin, J.N., Muhamad, M., Kiting, R., Yusuf, R.: Digital technology adoption for instruction aids: insight into teaching material content. In: Intelligent Strategies for ICT: Proceedings of ICTCS 2023, vol. 2. Springer, Cham (2024)
74. Alwi, A.S.Q., Ibrahim, R.: Isu terhadap Penggunaan Teknologi Media Digital dalam kalangan guru pelatih jurusan Pendidikan Khas. Firdaus J. **2**(2), 88–93 (2022). https://doi.org/10.37134/firdaus.vol2.2.9.2022
75. Zamri, F.A., Muhamad, N., Huda, M., Hashim, A.: Social media adoption for digital learning innovation: insights into building learning support. In: Cyber Security and Applications: Proceedings of ICTCS 2023, vol. 3. Springer, Cham (2024)
76. Huda, M., et al.: Trust in electronic record management system: insights from Islamic-based professional and moral engagement-based digital archive. In: Software Engineering Methods in Systems and Network Systems - Proceedings of 7th Computational Methods in Systems and Software 2023. Springer, Cham (2024)
77. Yahya, S.F., Othman, M.A.: Penggunaan video dalam Pengajaran dan Pembelajaran Pendidikan Moral Tingkatan 2. Firdaus J. **2**(2), 94–105 (2022). https://doi.org/10.37134/firdaus.vol2.2.10.2022
78. Rahman, M.H.A., Jaafar, J., Huda, M.: Information and communication skills for higher learners competence model. In: Software Engineering Methods in Systems and Network Systems - Proceedings of 7th Computational Methods in Systems and Software 2023. Springer, Cham (2024)
79. Cita Sari, D., et al.: Transformation of artificial intelligence in Islamic Edu with Ulul Albab Value (Global Challenge Perespective). Firdaus J. **3**(1), 1–9 (2023). https://doi.org/10.37134/firdaus.vol3.1.1.2023
80. Huda, M., et al.: From digital ethics to digital community: an Islamic principle on strengthening safety strategy on information. In: Software Engineering Methods in Systems and Network Systems - Proceedings of 7th Computational Methods in Systems and Software 2023. Springer, Cham (2024)
81. Zamri, F.A., Muhamad, N., Huda, M.: Information and communication technology skills for instruction performance: beliefs and experiences from public school educators. In: Software Engineering Methods in Systems and Network Systems - Proceedings of 7th Computational Methods in Systems and Software 2023. Springer, Cham (2024)
82. Masud, A., Borham, A.H., Huda, M., Rahim, M.M.A., Husain, H.: Managing information quality for learning instruction: Insights from public administration officers' experiences and practices. In: Software Engineering Methods in Systems and Network Systems - Proceedings of 7th Computational Methods in Systems and Software 2023. Springer, Cham (2024)
83. Huda, M., et al.: Digital record management in Islamic education institution: current trends on enhancing process and effectiveness through learning technology. In: Software Engineering Methods in Systems and Network Systems - Proceedings of 7th Computational Methods in Systems and Software 2023. Springer, Cham (2024)
84. Mohd Nawi, M.Z.: Media variations in education in Malaysia: a 21st century paradigm. Firdaus J. **3**(1), 77–95 (2023). https://doi.org/10.37134/firdaus.vol3.1.8.2023
85. Huda, M., et al.: From technology adaptation to technology adoption: an insight into public school administrative management. In: Proceedings of Ninth International Congress on Information and Communication Technology. ICICT 2024. Lecture Notes in Networks and Systems. Springer, Singapore (2024)
86. Tan, A.A., Huda, M., Rohim, M.A., Hassan, T.R.R., Ismail, A.: Chat GPT in supporting education instruction sector: an empirical literature review. In: Proceedings of Ninth International Congress on Information and Communication Technology. ICICT 2024. Lecture Notes in Networks and Systems. Springer, Singapore (2024)

87. Wahid, A., et al.: Augmented reality model in supporting instruction process: a critical review. In: Proceedings of Ninth International Congress on Information and Communication Technology. ICICT 2024. Lecture Notes in Networks and Systems. Springer, Singapore (2024)
88. Musolin, M.H., Serour, R.O.H., Hamid, S.A., Ismail, A., Huda, M., Rohim, M.A.: Developing personalized Islamic learning in digital age: pedagogical and technological integration for open learning resources (OLR). In: Proceedings of Ninth International Congress on Information and Communication Technology. ICICT 2024. Lecture Notes in Networks and Systems. Springer, Singapore (2024)
89. Musolin, M.H., Ismail, M.H., Huda, M., Hassan, T.R.R., Ismail, A.: Towards an Islamic education administration system: a critical contribution from technology adoption. In: Proceedings of Ninth International Congress on Information and Communication Technology. ICICT 2024. Lecture Notes in Networks and Systems. Springer, Singapore (2024)
90. Musolin, M.H., Ismail, M.H., Farhan, M., Rois, N., Huda, M., Rohim, M.A.: Understanding of artificial intelligence for Islamic education support and service: insights from empirical literature review. In: Proceedings of Ninth International Congress on Information and Communication Technology. ICICT 2024. Lecture Notes in Networks and Systems. Springer, Singapore (2024)
91. Susilowati, T., et al.: Getting parents involved in child's school: using attendance application system based on SMS gateway. Int. J. Eng. Technol. **7**(2.27), 167–174 (2018)
92. Latif, M.K., Md Saad, R., Abd Hamid, S.: Islamic education teachers' competency in teaching Qiraat Sab'ah for the Quranic Class. Firdaus J. **3**(1), 19–27 (2023). https://doi.org/10.37134/firdaus.vol3.1.3.2023
93. Susilowati, T., et al.: Learning application of Lampung language based on multimedia software. Int. J. Eng. Technol. **7**(2.27), 175–181 (2018)
94. Abadi, S., et al.: Application model of k-means clustering: insights into promotion strategy of vocational high school. Int. J. Eng. Technol. **7**(2.27), 182–187 (2018)
95. Huda, M., et al.: Building harmony in diverse society: insights from practical wisdom. Int. J. Ethics Syst. (2020). https://doi.org/10.1108/IJOES-11-2017-0208
96. Aminudin, N., et al.: The family hope program using AHP method. Int. J. Eng. Technol. **7**(2.27), 188–193 (2018)
97. Wulandari, et al.: Design of library application system. Int. J. Eng. Technol. **7**(2.27), 199–204 (2018)
98. Aminudin, N., et al.: Higher education selection using simple additive weighting. Int. J. Eng. Technol. **7**(2.27), 211–217 (2018)
99. Zainuri, A., Huda, M.: Empowering cooperative teamwork for community service sustainability: insights from service learning. Sustainability **15**(5), 4551 (2023)
100. Maseleno, A., et al.: Hau-Kashyap approach for student's level of expertise. Egypt. Inform. J. **20**(1), 27–32 (2019)
101. Richey, R., Klein, J.: Design and Development Research: Methods, Strategies and Issues. Lawrence Erlbaum Associates, Mahwah (2007)
102. Huda, M.: Empowering application strategy in the technology adoption: insights from professional and ethical engagement. J. Sci. Technol. Policy Manag. **10**(1), 172–192 (2019)
103. Yousefi, S., Tosarkani, B.M.: Exploring the role of blockchain technology in improving sustainable supply chain performance: a system-analysis-based approach. IEEE Trans. Eng. Manag. (2023)
104. Kembauw, E., Soekiman, J.F.X.S.E., Lydia, L., Shankar, K., Huda, M.: Benefits of corporate mentoring for business organization. J. Crit. Rev. **6**(5), 101–106 (2019)
105. Huda, M.: Towards digital access during pandemic age: better learning service or adaptation struggling? Foresight **25**(1), 82–107 (2023). https://doi.org/10.1108/FS-09-2021-0184

106. Huda, M.: Digital marketplace for tourism resilience in the pandemic age: voices from budget hotel customers. Int. J. Organ. Anal. **31**(1), 149–167 (2023). https://doi.org/10.1108/IJOA-10-2021-2987
107. Rosa, A.T.R., Pustokhina, I.V., Lydia, E.L., Shankar, K., Huda, M.: Concept of electronic document management system (EDMS) as an efficient tool for storing document. J. Crit. Rev. **6**(5), 85–90 (2019)
108. Huda, M., et al.: Strategic role of trust in digital communication: critical insights into building organizational sustainability. In: Arai, K. (ed.) Proceedings of the Future Technologies Conference (FTC). FTC 2023. LNNS, vol. 815, pp. 387–403. Springer, Cham (2023)
109. Abdul Aziz, N.A., Mohd Razali, F., Saari, C.Z.: Penggunaaan Media Sosial dari Perspektif Psiko Spiritual Islam. Firdaus J. **2**(1), 65–75 (2022). https://doi.org/10.37134/firdaus.vol2.1.6.2022
110. Hanafi, H.F., et al.: A review of learner's model for programming in teaching and learning. J. Adv. Res. Appl. Sci. Eng. Technol. **33**(3), 169–184 (2024)
111. Huda, M., Bakar, A.: Culturally responsive and communicative teaching for multicultural integration: qualitative analysis from public secondary school. Qual. Res. J. (2024, aheadofprint). https://doi.org/10.1108/QRJ-07-2023-0123
112. Syafri, N., Ali, A.H., Ramli, S.: Penggunaan kaedah Inovasi Sambung dan Baca Bahasa Arab (SaBBAr) dalam meningkatkan kemahiran membaca perkataan Bahasa Arab murid di Sekolah Rendah. Firdaus J. **2**(2), 62–71 (2022). https://doi.org/10.37134/firdaus.vol2.2.6.2022
113. Huda, M., et al.: Enhancing digital leadership direction: insight into empowering gender violence prevention. In: Mishra, D., et al. (eds.) Communication Technology and Gender Violence. Springer, Cham (2024)
114. Ab Rahim, N.M.Z., Saari, Z., Mohamad, A.M., Rashid, M.H., Mohamad Norzilan, N.I.: Konsep Ulul Albab dalam Al-Quran dan hubungannya dengan pembelajaran kursus Sains, Teknologi dan Manusia di UTM Kuala Lumpur. Firdaus J. **2**(2), 72–78 (2022). https://doi.org/10.37134/firdaus.vol2.2.7.2022
115. Huda, M., et al.: Understanding of digital ethics for information trust: a critical insight into gender violence anticipation. In: Mishra, D., et al. (eds.) Communication Technology and Gender Violence. Springer, Cham (2024)
116. Hanafi, H.F., Wahab, M.H.A., Selamat, A.Z., Masnan, A.H., Huda, M.A.: Systematic review of augmented reality in multimedia learning outcomes in education. In: Proceedings of the 12th International Conference on Intelligent Human Computer Interaction, IHCI 2020, Part II 12, Daegu, South Korea, 24–26 November 2020, pp. 63–72. Springer, Cham (2021)
117. Huda, M., Gusmian, I., Mulyo, M.T.: Towards eco-friendly responsibilities: Indonesia field school model cross review. J. Comp. Asian Dev. (JCAD) **18**(2), 1–12 (2021)
118. Leh, F.C., Anduroh, A., Huda, M.: Level of knowledge, skills and attitude of trainee teachers on Web 2.0 applications in teaching geography in Malaysia schools. Heliyon **7**(12) (2021)
119. Sukadari, S., Huda, M.: Culture sustainability through co-curricular learning program: learning Batik cross review. Educ. Sci. **11**(11), 736 (2021)
120. Rachman, A., Oktoviani, I., Manurung, P.: Digitization of Zakat and charity BAZNAS Tangerang City through crowdfunding platform tangerangsedekah. id. Firdaus J. **3**(1), 96–106 (2023). https://doi.org/10.37134/firdaus.vol3.1.9.2023
121. Huda, M., Kartanegara, M.: Islamic spiritual character values of al-Zarnūjī's Ta 'līm al-Muta 'allim. Mediterranean J. Soc. Sci. **6**(4S2), 229–235 (2015)
122. Huda, M., Yusuf, J.B., Jasmi, K.A., Nasir, G.A.: Understanding comprehensive learning requirements in the light of al-Zarnūjī's Ta'līm al-Muta'allim. SAGE Open **6**(4), 1–14 (2016)
123. Huda, M., Yusuf, J.B., Jasmi, K.A., Zakaria, G.N.: Al-Zarnūjī's concept of knowledge ('ilm). SAGE Open **6**(3), 1–13 (2016)

124. Huda, M., Jasmi, K.A., Mohamed, A.K., Wan Embong, W.H., Safar, J.: Philosophical investigation of Al-Zarnuji's Ta'lim al-Muta'allim: strengthening ethical engagement into teaching and learning. Soc. Sci. **11**(22), 5516–5551 (2016)
125. Huda, M., et al.: Innovative teaching in higher education: the big data approach. Turk. Online J. Educ. Technol. **15**(Spec. Issue), 1210–1216 (2016)
126. Othman, R., Shahrill, M., Mundia, L., Tan, A., Huda, M.: Investigating the relationship between the student's ability and learning preferences: evidence from year 7 mathematics students. New Educ. Rev. **44**(2), 125–138 (2016)
127. Huda, M., Sabani, N., Shahrill, M., Jasmi, K. A., Basiron, B., Mustari, M.I.: Empowering learning culture as student identity construction in higher education. In: Shahriar, A., Syed, G. (eds.) Student Culture and Identity in Higher Education, pp. 160–179. IGI Global, Hershey (2017). https://doi.org/10.4018/978-1-5225-2551-6.ch010
128. Huda, M., et al.: Empowering children with adaptive technology skills: careful engagement in the digital information age. Int. Electron. J. Element. Educ. **9**(3), 693–708 (2017)
129. Huda, M., Jasmi, K.A., Basiran, B., Mustari, M.I.B., Sabani, A.N.: Traditional wisdom on sustainable learning: an insightful view from Al-Zarnuji's Ta 'lim al-Muta 'allim. SAGE Open **7**(1), 1–8 (2017)
130. Huda, M., Jasmi, K.A., Alas, Y., Qodriah, S.L., Dacholfany, M.I., Jamsari, E.A.: Empowering civic responsibility: insights from service learning. In: Burton, S. (ed.) Engaged Scholarship and Civic Responsibility in Higher Education, pp. 144–165. IGI Global, Hershey (2017). https://doi.org/10.4018/978-1-5225-3649-9.ch007
131. Huda, M., et al.: Innovative e-therapy service in higher education: mobile application design. Int. J. Interact. Mob. Technol. **11**(4), 83–94 (2017)
132. Huda, M., et al.: From live interaction to virtual interaction: an exposure on the moral engagement in the digital era. J. Theor. Appl. Inf. Technol. **95**(19), 4964–4972 (2017)
133. Huda, M., Maseleno, A., Jasmi, K.A., Mustari, I., Basiron, B.: Strengthening interaction from direct to virtual basis: insights from ethical and professional empowerment. Int. J. Appl. Eng. Res. **12**(17), 6901–6909 (2017)
134. Huda, M., Haron, Z., Ripin, M.N., Hehsan, A., Yaacob, A.B.C.: Exploring innovative learning environment (ILE): big data era. Int. J. Appl. Eng. Res. **12**(17), 6678–6685 (2017)
135. Maseleno, A., et al.: Combining the previous measure of evidence to educational entrance examination. J. Artif. Intell. **10**(3), 85–90 (2017)

From Digital Ethics to Digital Community: An Islamic Principle on Strengthening Safety Strategy on Information

Miftachul Huda[1(✉)], Mukhamad Hadi Musolin[2], Mohamad Hazli Ismail[2], Andi Muhammad Yauri[3], Abu Bakar[3], Muhammad Zuhri[3], Mujahidin[3], and Uswatun Hasanah[3]

[1] Sultan Idris Education University, Tanjung Malim, Malaysia
halimelhuda@gmail.com
[2] Sultan Abdul Halim Mu'adzam Shah International Islamic University, Kuala Ketil, Malaysia
[3] State Institute for Islamic Studies, Bone, South Sulawesi, Indonesia

Abstract. This paper aims to examine the demanding needs of digital ethics elaborated as a strategic foundation to expand safety concern amongst digital community and partnership. The literature review was applied from the referred articles from peer-reviewed journals, proceedings, chapter and books related to the topic. The finding reveals that the significant alignment of digital ethics in expanding the digital partnership skills has a core value to expand the digital community. This paper may contribute to give insights into developing the outstanding interplay of digital ethics as a safety strategy to achieve digital community and partnership.

Keywords: digital ethics · digital partnership skills · digital community · safety strategy · Islamic critical insight

1 Introduction

In the recent years, the advanced technology has been widely transformed into the total life of human society through many sectors, such as running business, education process, and social interaction platform and entertainment purposes. On this view, the expansion of sophisticated features has led to the supporting pathway through giving the services to the society [1, 39]. With such advancement, the potentials of giving the ultimate benefit to the human could also create the challenging issues, like untrusted information, which requires the ethical engagement on solving it properly. Concerning the need to enhance the pathway of ethics in the context of digital practices should do with reactualising the digital ethics [40, 41]. The digital ethics with its features in constructing the digital environment responsibility requires us to have a sufficient application process and procedure in accessing and adapting technology [2, 42]. Moreover, the technology accomplishment in bringing the latest development articulated into the features has performed in helping the society.

In addition, the emerging potentials of recent development of technology has called for the digital ethics empowerment in bringing the widely sufficient comprehension to

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 165–182, 2024.
https://doi.org/10.1007/978-3-031-53552-9_15

look into detail about the need of ethical engagement apart from professional enhancement [3, 43]. At this point of view, the digital environment scenario needs to enhance the strategic empowerment to transfer the responsibility awareness towards the electronic portable devices, online interaction and cyber space activity related to the human society. In particular, the scenario on enhancing the strategic practice allowed the digital ethics on monitoring the process of technologically organized morality [44, 45]. The first point, which needs to concern, is related to the comprehension pathway in the sense that knowledge expertise to drive the users' technology practice with the proper manner [4]. On this view, the digital ethics is required to possess the sufficiency of obtaining the understanding towards the digitally social interaction and communication [46, 47]. As a result, the way of ethical practice with appropriate manner refers to occupy the value in giving feedback to strengthen capacity and capability.

With this regard, the need to have a sufficient description of digital ethics should be taken into consideration in raising the society's awareness. There are many studies on the way on how to expand the ethics on digital technology in giving the contribution to the society at any sector [5, 48]. As such, this paper aims to examine the demanding needs of digital ethics elaborated as a strategic foundation to expand safety concern amongst digital community and partnership. In order to achieve this scenario, the required practice on searching for the literature should be made in reviewing process from the peer-reviewed and referred articles from journals, proceedings, chapter and books related to the topic. The potential finding showed the significances of aligning the digital ethics to exposure the digital partnership skills to monitor the users' society in the technology adoption. In particular, the ultimate point of this study refers to give the value in expanding the digital ethics to sustain the digital community through enhancing the simultaneous interplay on adapting technology with ethics to drive into creating the safety amidst digital community and partnership.

2 Literature Review

2.1 Understanding of Digital Ethics

Referring to the ethical standard in guiding the digital practices, the main component of digital ethics as the pathway to drive into bringing the guideline of transforming the digital activities aims to ensure the process and procedure in the digital society with the safety concern [6, 49]. With this regard, the digital ethics comprehension needs to expand the expertise to respond and reserve what the point is on achieving the digital practice to digital sustainability. The need to have a transformation in driving the direction to the society on the online practice should bring the digital ethics translated into the contextual circumstance [50, 51]. On this view, the extent of digital expansion should come up with directing the society amidst the millennial age transformation into technology expert and concern. It indicated that the proper arrangement refers to enhance the digital ethics empowerment in the ethical and moral balance [7, 52]. At this point of view, the mutual concern to contribute in transforming the digital society by creating the organizing strategy in ensuring the information data with safety and security.

In addition, the concern of digital ethics in helping to transform the way of the information use in enhancing the digital society awareness needs to focus on expanding the

ethical concern to drive the pathway of direction amidst the digital environment. Moreover, the digital ethics with its significant contribution to expand governing the ethical engagement on underlying the digital practice guideline should come up with building the strategic consciousness towards information dissemination process and procedure [8, 53]. With this regard, the strategic engagement in helping to disseminate the information across the platform of online into sharing with others' membership. As a result, the digital ethics needs to have paid a particular attention on strategizing to enhance the digital awareness amidst the cyber society [54, 55]. On this view, attempts to bring the ethical principles in giving insights into the digital practice would be the strategic achievement to enhance the digital ethics. As a result, the strategic enhancement on maintaining the proper manner on considering the digital ethics refers to contribute in providing the essentials of comprehending and practicing the guideline [9, 56]. As such, the expertise to maintain in giving the line in guiding the digital practices should bring along with building the ethical engagement as the standing point of arranging the entire process.

In line with concerning the digital ethics as the example on monitoring the cyber and online practices, the guideway on managing the appropriate manner in enhancing the standard of social networking site (SNS) for instance needs to concern in arranging the influential value to underlie the ethics. With this regard, the potential value in increasing the digital practice requires to provide the strategic awareness of digital ethics application [57, 58]. Through playing a significant role in reaching to give on covering the entire online process, the mutual requirement of digital ethics' user awareness should bring along with possessing the comprehension in giving the direction pathway into the proper manner arrangement [10, 59]. On this view, the sufficient understanding of managing the digital practices with the ethical standard concern needs to arrange the way of occupying the prominent achievement with directing the digital ethics. At this point of view, the clear arrangement on obtaining the potentials of giving the beneficial value is required to enhance the personal and social capacity and capability with the digital ethics [60, 61]. As a result, the insightful value to strengthen the individual arrangement to govern in enhancing the strategic effort to raise the responsibility on digital ethics should do with taking the particular consideration to empower the digital practices and environment [11, 62]. As such, the digital users played a role in ensuring the digital environment with its potential importance in taking the strategic pathway to perform the sufficient facilities to adapt and adopt the proper use of digital engagement with the ethical manner conviction.

2.2 Digital Ethics as Strategic Skills in Digital Community

The consistency of arranging the digital ethics is concisely expanded to continue the strategic arrangement of managing the digital community in strengthening the partnership amidst the cyber society. At this point of view, the strategic pathway in driving the digital ethics in underlying the digital practice refers to enhance the committed awareness on online practice [12, 63]. Moreover, the digital practice is supposedly engaged to have on applying for combining the digital partnership skills in digital environment. Amidst the online environment, the particular arrangement of consolidating the digital ethics empowerment needs to have the constant commitment in continuing the online practice [64, 65]. Moreover, the need to incorporate the ethical consistency in leading

to give a constant moral standard should commit with giving insights into contributing it into the digital society [13, 66]. As a result, the committed awareness in gaining the digital partnership should do with expanding to the digital practice arrangement amidst the online environment. In particular, the wider sphere of multiple purposes in achieving the knowledge understanding refers to give a valuable point of moral stability [67, 68]. Through the comprehension incorporated with the potentials of ethical engagement, the particular point is supposedly to build the digital partnership skills arranged into the attainment of stabilizing the digital ethics' main concern. The arrangement on making real into the digital circumstance could be well integrated in the human's inner pathway [69, 70].

In line with obtaining the digital ethics arrangement, the strategic governance on driving the ethical standard commencement would be standing point on the digital partnership in digital environment. With this regard, it is necessary to point out the value of its particular part of digital ethics in carrying out the strategic commitment to apply for the digital partnership [14, 71]. It is necessary to help making concise of giving the value through a sufficient comprehension followed by the digital skills enhancement. On this view, the balance between knowledge and skills in bringing into the active incorporation needs to maintain the digital society in driving the technology use properly in line with the context. As a result, the need to take the knowledge attribution of digital circumstance in bringing together with the comprehension and practical abilities needs to consider both skills and knowledge, which would be useful to drive in the online practices [15, 72]. As such, the wider context of digital society to integrate within the digital activities aims to enhance the careful consideration to commit in the digital partnership and skills.

In further, the important part on monitoring the digital ethics in underlying the partnership requires the comprehensive engagement to maintain the privacy and security concern. As the significant contribution in building the digital skills, the concise arrangement on committing with the particular engagement to give the freethinking is important to play a significant role to be actively managed into digital ethics commitment [16, 73]. With this regard, the strategic attempts on reenergizing the digital ethics could give insights into valuing the basis of the actual reality to strengthen the digital skills into the better service. Moreover, the balance between personal data arrangement and digital partnership skills refers to point out disseminating the responsible awareness to be aware of the consequence caused by the digital environment [17, 74]. As a result, it has to do with possessing the clear purpose on the multiple tasks to lead to an active digital partnership so that it can provide the proper service as incorporated with the digital ethics. As such, the emerging potential risks in digital practices aims at enhancing the potentially digital partnership incorporated with the sufficient comprehension and the actualization [75, 76]. Thus, the maintenance process is supposedly to give insightful value in guiding the cyber arrangement for online practices to be integrated to have actualized within the appropriate way.

2.3 Digital Ethics and Digital Trust as Key Element for Safety Strategy

Considered as the conceptualizing process with bringing the clear information from the sources, the essence of digital trust yields to become the key direction to drive the digital

activities across the varied purposes [18, 77]. The process of disseminating information and data from the supplier to the receiver amidst the digital environment, for instance, requires the serious concern in maintaining the clarity, consequence and stability. With this regard, the urgent points of having the digital ethics could be the significant role in transmitting the foundational value to arrange the inner pathway of digital users. The important pathway on giving the value of digital ethics refers to enable the users in maintaining to monitoring the progression of continuing online environment. The strategic attempts in continuing the value of digital ethics on digital partnership skills in enabling the digital users to continue their activities amidst the online environment [19, 78]. The key concept of digital ethics with its significance aims to advance the trust-committed digital partnership amidst the global communities. On this view, the digital concept needs to integrate in allocating the digital users to ensure the information and personal data in achieving the safety concern.

In addition, the strategic enhancement to achieve the privacy should do with in reaching to the safety concern. The particular point on enhancing the particular way to ensure the sufficient detail of knowledge needs to consider in having the comprehension together with adaptive practice [20, 79]. In terms of actualizing the digital practices, the attempts to integrate the digital ethics with its significant value pointed out building the digital society in bringing the active skills. With this regard, the key element in advancing the digital ethics needs to pay attention on being aware of trust assurance through directing the pathway in achieving the inner pathway of digital user. As a result, the strategic enhancement to expand the commitment on committing with the clear ethical standards is supposed to govern an extensive portrait to look into reflecting the partnership skills amidst the digital society [21, 80]. Moreover, the process on advancing the active participation in digital society should be looking into detail about what to do in line with underlying the digital users. The sufficient contribution with having the careful engagement strategy should be concisely reflected with dealing the strategic process in resulting to the digital partnership [81, 82].

In further, the digital ethics on possessing the pathway to drive in governing the process in the online environment within the appropriate manner. In order to have the attainment to provide the significant contribution to have correlated to the arrangement of proper way, attempts to ensure the digital practices within a smooth pathway should be pointed out having an active engagement to continue serving in treating the digital users with the given services [22, 83]. The strategic attempts to provide the allocation stability from both senders to receiver requires having the digital partnership in contributing to give the value in forwarding the possible feedback. On this view, the additional point on enhancing the significance of the consistency in committing with the digital ethics refers to consistently in building the crucial value of digital ethics [84, 85]. As a result, it aims to enhance the careful engagement on upholding the extensive empowerment of ethical manner standards [86, 87]. In order to obtain the significant contribution of active support on carrying out the digital practice, the strategic assessment on earning the mutual support to achieve the digital trust is required to have the sufficient adaptation on driving the technology use.

3 Methodology

With having the critical review from the peer-reviewed articles, the aim of this study was to examine the empowerment strategy of digital ethics to digital community. The main point of having the literature analysis aims at critically exploring the application procedure to strengthen the safety concern. Through providing the strategic awareness of digital ethics on information, attempts to empower strategically professional and ethical balance are assessed to organise in examining the recent reviews on the related issues. Moreover, the quality awareness on managing the information data in achieving the safety assurance is required to have a sufficient assessment on investigating the content and context of digital ethics to digital community. As a result, the attempts on empowering the safety strategy on data information from peer-reviewed articles from journals, chapters, proceeding and books in particular are made through the descriptive analysis. The significant phases include integrating, evaluating and interpreting the findings from the varied types of employed research, which came from the recent works of grounded theory.

4 Analysis and Discussion

4.1 Expanding Safety Strategy Through Digital Ethics for Information Assurance

In the attempts on continuing the safety strategy, the adaptive practice of digital ethics is being one of the important aspects to drive the key performance of cyber oriented user practices. Through the digital ethics-based safety strategy, the particular point on monitoring the digital practices aims at building the substantive key point on strategic partnership in digital environment [23, 88]. The particular result is given into empowering the strategic approach to consider the digital ethics in reflecting the significant point on elaborating the digital practices. The digital skills in elaborating the proper association to fulfill the digital revolution industry age comes from an emerging technology enhancement. Such arrangement on digital ethics' arrangement is needed to likely fulfil the potentials of good attempts on obtaining the acceptance through trust [89, 90]. The value of driving digital ethics is potentially enhanced with bringing the safety strategy in the application design in the digital community context. In particular, the strategic chances on achieving the target on for safety achievement requires following the distinctive characteristics on digital practices with having a fair consistency [24, 91]. The extent of safety strategy as the main concern on gaining the active engagement in the digital environment should do with further exploring the trust as key elements of inner pathway to enhance the foundation.

In addition, the advancement of digital trust assurance is important to bridge in enhancing the accuracy of information being the outstanding value in driving the digital practice. As a result, the attempts to obtain in gaining the digital ethics should bring along with developing the partnership engagement skills [25, 92]. The strategic pathway of incorporating the digital ethics in accordance with running the artificial intelligence aims at designing the digital society as the users in adhering to the trust commitment. On this view, the basis of instructional design refers to give an insight into actualizing the

digital society through enabling the strategic way to drive into energizing the privacy and security concern [93, 94]. As a result, distinctive trust is being the necessary act to help improving the digital skills to take into considering the particular way on managing the digital society in reaching to the advance the safety strategy [95, 96]. In order to achieve such arrangement, the particular point in advancing the safety achievement refers to provide the mutual line in building the mutual engagement between convenience and efficiency [26, 97]. In particular, the digital society context should do with continuing the commitment to bridge the digital trust in achieving both safety and feel sufficiency. As such, the comfortable aspects are compassionately elaborated in helping the society to possess the progressive accountability amidst the digital practices [98, 99]. The particular result in developing the strategic point refers to help contribute in disseminating the important value as the elements of digital skills.

In further, the essence of digital ethics combined with the skills would expand in accessing the accommodation in committing with the belief and continuity. It is necessary to help increase the digital trust in giving the significant value to achieve the transparency in information context [27, 100]. The consistency and commitment are being the particular orientation on contributing the digital ethics actualized to maintain the strategic rules in term of application amidst the digital circumstance. With this regard, attempts to organize the continued practice in the digital society should bring along with governing the strategic purpose on digital environment expansion [101, 102]. On this view, the particular attention needs to take consider in managing the way to ensure the safety concern. Such arrangement refers to the trust and transparency, which needs to be collaborated with the application rules to further forward in the implementation context amidst the digital organization [28, 103]. At this point of view, the further action on reenergizing the practical use of digital ethics is being the strategic approach to put forward the succession in ensuring the digital practices with the security and privacy concern. Moreover, the critical exploration points out taking precedence over the safety succession in underlying the online practices [103, 104]. Transformed into such sector like education technology, business strategy, and societal purposes, the continued attempts on embedding the general public sphere refers to drive the direction on driving the digital ethics' commitment to enhance the strategic awareness to expand the mutual trust.

4.2 Enhancing Safety Strategy Through Digital Ethics-Based Information Transparency Assurance

The strategic enhancement to achieve the information clarity and transparency yields to empower the attempts in expanding the digital ethics and skills orientation [29]. Moreover, the digital ethics management in underlying the online practices is concisely set up with the digital framework engagement [106, 107]. In particular, the further expansion on building a mutual understanding and code of managing the future prospect needs to strategize the digital ethics required to integrate the ethical standard. On this view, the particular attention should be given in resulting the sufficient recognizance to determine the digital practices together with ethical manner as the systematic norms to transmit in achieving the information transparency [30, 108]. As a result, the further direction to commit with the ethical manner is actualized to empower the strategic application

use in implementing the digital technologies. In particular, the digital ethics framework requires recognizing the assimilation to sustain the outcome in fostering the information transparency. In the attempts to generate the organizations, the principals of digital ethics should be committed in order to take the beneficial value in adhering the interconnection of digital users and services [109, 110]. Empowered with the commitment to continuously engaging the norms to build the consistency amongst the users, achieving the information transparency is being the necessary part in ensuring the incorporation process for the digital ethics in transmitting and transforming into the environmental basis.

In addition, the consistency on managing the digital users refers to strategize in incorporating the digital ethics in ensuring the information transparency [31, 111]. It is mainly during the transmission process that the careful engagement on adapting the moral standard played a significant role in achieving the transparency. On this view, the attempts to start with the inner commitment continued to practice in reenergizing the digital environment to reflect the online practices [32, 112]. With this regard, the requirement is that ensuring the digital ethics awareness with its important consistency refers to take note in empowering the digital environment awareness. In order to have a transparency assurance, it is required mainly during the actual digital practice in order to cooperate with knowledge comprehension and skill empowerment. As such, the active engagement in considering the succession on moral principles-digital ethics would be the standing point in governing the sufficient transparency [113, 114]. Attempts to empower the digital practices should strengthen to resulting in strategizing the ethical codes in driving the direction pathway among the digital users' manner.

In further, the particular way to look into detail about the moral and professional standard would lead to enhance the stability to continue the digital ethics with the digital environment manner reflection. Committing with a clear comprehension to contribute the actual manifestation refers to possess the understanding pathway of digital ethics in order to maintain the clear information and accuracy [33, 115]. On this view, the clear comprehension about the digital practices with its strategic application and use yields the orientation on expanding the digital environment establishment. With this regard, the digitally planned arrangement in underlying the basis of digital ethics points out the strategy to achieve professional and ethical engagement [116, 117]. The achievement procedure as the result on monitoring process to have the sufficient look into digital adaptability and acceptability requires the further application to concisely bring the digital partnership [32, 116]. Moreover, considering the ethical standard dimension in determining the further accommodation for instance refers to apply for the emphatically governed norms to deliver in translating into the digital environment. Such orientation needs to organize the proper direction adapted to enhance the digital ethics in the organization management.

4.3 Empowering Safety Strategy Through Digital Ethics on Information Accuracy

The strategic pathway to ensure the clear and clarity of information could start with upholding the strategic approach to create the habitual action with digital trust. Amidst the digital society, attempts to take a consideration on adapting the proper arrangement

in the workplace are clearly emphasized to enhance the acceptable manner reflected in digital ethics [34, 118]. In the attempts to expand the digital practices, having the sufficient comprehension should point out taking the adaptive empowerment to rebuild the digital ethics' commitment. With this regard, the strategic approach to generate the digital environment should govern in building the initial arrangement followed by getting well-prepared norms to drive a key direction on the information accuracy achievement. Moreover, the continued practice to stabilize the accuracy enhancement requires disseminating the strategic features on underlying the process to ensure the foundation amongst the digital users [35, 119]. Managing the proper manner in building the digital society refers to point out the significance of trust engagement to consider in enhancing the safety orientation. The further consideration with the particular way to coordinate the digital ethics should commit in considering the inner pathway of digital users [120, 121]. As a result, attempts to achieve the digital trust with bringing its featured characteristics are strategically taken a beneficial value to drive the online users. Amidst the digital society, taking benefit of the digital practice is being the fundamental point to enhance the commitment of inner pathway among the digital users in order to achieve the digital trust [122, 123]. On this view, achieving the digital trust with its featured characteristics amidst the digital society needs to continue strategizing the digital online environment.

In addition, taking the beneficial value on driving the digital technology improvement to result in expanding the digital users should govern in integrating the digital ethics with its ultimate point on applying for the mutual direction amidst the digital complexities. With this regard, the comprehensive compliment to admit in empowering the digital ethics' complete arrangement refers to positively continue in expanding the strategic empowerment to enhance the stability [36, 124]. Moreover, the strategic accomplishment as an attempt in building the digital ethics' commitment needs to drive the digital users in underlying the digital technology application and use amidst the online circumstance. In particular, the mutual involvement on behaving digital skills and ethics empowerment continued to be an important part to help developing the further assessment to ensure the information accuracy [36, 125]. In order to advance the appreciation on the recent advancement on digital ethics, the strategic attempts in empowering the commitment to transmit the awareness on digital ethics needs to engage actively into information process in building the sufficient transparency [126, 127]. As a result, the further expansion on building both clarity and transparency of information is required to the digital users to allow performing the continuity and consistency through the clarification procedure towards the received information [37, 128]. At this point of view, the important part on enhancing the digital ethics' commitment needs to giver in ensuring the digital norms in contributing the digital user to devote into the self-sufficiency amidst the digital circumstance.

In line with enhancing the digital skills to drive the partnership direction amidst the digital society, attempts to possess the sufficient comprehension through taking consideration about the application and rule on digital practice are strategically enhanced to follow the significance of media literacy education [129, 130]. At this point of view, the distinctive features of digital ethics with its particularly logical view should bring along with looking into detail on concerning the mutual relationships among the digital users. With this regard, the particular engagement on comprehending the digital ethics

translated into the link of both corporations and clients for instance needs to arrange in possessing the level of individuals and socials with continually applying for the digital ethics-commitment [38]. On this view, ensuring the information accuracy requires to build the consistency empowerment on stabilizing the continuity in digital interaction amidst the digital society environment [132, 133]. In particular, the further exploration on grabbing both knowledge and adaptive practice refers to obtain the mutual line to carry out the digital interaction with bringing the essence of digital ethics. Moreover, the adaptive enhancement to look into detail about behaving the communication skills needs to integrate the digital ethics commitment to empower as the pattern style of digital users [134, 135]. It is important to note that the basis of ethical foundation to underlie the digital practice in resulting in accommodating the strategic corporations followed with the communication approach in online basis with the digital ethics empowerment.

5 Conclusion

This paper aims to examine the demanding needs of digital ethics elaborated as a strategic foundation to expand safety concern amongst digital community and partnership. The literature review was applied from the referred articles from peer-reviewed journals, proceedings, chapter and books related to the topic. The finding reveals that the significant alignment of digital ethics in expanding the digital partnership skills has a core value to expand the digital community. This paper may contribute to give insights into developing the outstanding interplay of digital ethics as a safety strategy to achieve digital community and partnership. The potential finding showed the significances of aligning the digital ethics to exposure the digital partnership skills to monitor the users' society in the technology adoption. In particular, the ultimate point of this study refers to give the value in expanding the digital ethics to sustain the digital community through enhancing the simultaneous interplay on adapting technology with ethics to drive into creating the safety amidst digital community and partnership.

References

1. Abad-Segura, E., González-Zamar, M.D., Infante-Moro, J.C., Ruipérez Gar-cía, G.: Sustainable management of digital transformation in higher education: global research trends. Sustainability **12**(5), 2107 (2020)
2. Balyer, A., Öz, Ö.: Academicians' views on digital transformation in education. Int. Online J. Educ. Teach. **5**(4), 809–830 (2018)
3. Buchholz, B.A., DeHart, J., Moorman, G.: Digital citizenship during a global pandemic: moving beyond digital literacy. J. Adolesc. Health. **64**(1), 11–17 (2020)
4. Capurro, R.: Digital ethics. In: Global Forum on Civilization and Peace, pp. 207–216 (2009)
5. Falloon, G.: From digital literacy to digital competence: the teacher digital competency (TDC) framework. Educ. Tech. Res. Dev. **68**(5), 2449–2472 (2020). https://doi.org/10.1007/s11423-020-09767-4
6. Feerrar, J.: Development of a framework for digital literacy. Reference Services Review (2019)
7. Floridi, L., Cath, C., Taddeo, M.: Digital ethics: its nature and scope. In: Öhman, C., Watson, D. (eds.) The 2018 Yearbook of the Digital Ethics Lab. Digital Ethics Lab Yearbook, pp. 9−17. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-17152-0_2

8. Huda, M.: Empowering application strategy in the technology adoption: insights from professional and ethical engagement. J. Sci. Technol. Policy Manag. **10**(1), 172–192 (2019)
9. Huda, M.: Towards an adaptive ethics on social networking sites (SNS): a critical reflection. J. Inf. Commun. Ethics Soc. **20**(2), 273–290 (2022)
10. Huda, M., Hashim, A.: Towards professional and ethical balance: insights into application strategy on media literacy education. Kybernetes **51**(3), 1280–1300 (2022)
11. Juškevičienė, A., DagienĖ, V.: Computational thinking relationship with digital competence. Inform. Educ. **17**(2), 265–284 (2018)
12. Kateryna, A., Oleksandr, R., Mariia, T., Iryna, S., Evgen, K., Anastasiia, L.: Digital literacy development trends in the professional environment. Int. J. Learn. Teach. Educ. Res. **19**(7), 55–79 (2020)
13. Martzoukou, K., Fulton, C., Kostagiolas, P., Lavranos, C.: A study of higher education students' self-perceived digital competences for learning and everyday life online participation. J. Documentation (2020)
14. Nedungadi, P.P., Menon, R., Gutjahr, G., Erickson, L., Raman, R.: Towards an inclusive digital literacy framework for digital India. Educ.+ Training **60**(6), 516-528 (2018)
15. Passey, D., Shonfeld, M., Appleby, L., Judge, M., Saito, T., Smits, A.: Digital agency: empowering equity in and through education. Technol. Knowl. Learn. **23**(3), 425–439 (2018)
16. Porat, E., Blau, I., Barak, A.: Measuring digital literacies: junior high-school students' perceived competencies versus actual performance. Comput. Educ. **126**, 23–36 (2018)
17. Priyono, A., Moin, A., Putri, V.N.A.O.: Identifying digital transformation paths in the business model of SMEs during the COVID-19 pandemic. J. Open Innov. Technol. Mark. Complex. **6**(4), 104 (2020)
18. Reyman, J., Sparby, E.M.: Digital Ethics. Routledge, New York-London (2019)
19. Sánchez-Cruzado, C., Santiago Campión, R., Sánchez-Compaña, M.: Teacher digital literacy: the indisputable challenge after COVID-19. Sustainability **13**(4), 1858 (2021)
20. Sarbadhikari, S.N., Pradhan, K.B.: The need for developing technology-enabled, safe, and ethical workforce for healthcare delivery. Saf. Health Work **11**(4), 533–536 (2020)
21. Saura, J.R., Ribeiro-Soriano, D., Palacios-Marqués, D.: From user-generated data to data-driven innovation: a research agenda to understand user privacy in digital markets. Int. J. Inf. Manage. **60**, 102331 (2021)
22. Sheikh, A., et al.: Health information technology and digital innovation for national learning health and care systems. Lancet Digit. Health **3**(6), e383–e396 (2021)
23. Suwana, F.: Content, changers, community and collaboration: expanding digital media literacy initiatives. Media Pract. Educ. **22**(2), 153–170 (2021)
24. Torous, J., Myrick, K.J., Rauseo-Ricupero, N., Firth, J.: Digital mental health and COVID-19: using technology today to accelerate the curve on access and quality tomorrow. JMIR Mental Health **7**(3), e18848 (2020)
25. Zhao, Y., Llorente, A.M.P., Gómez, M.C.S.: Digital competence in higher education research: a systematic literature review. Comput. Educ. **168**, 104212 (2021)
26. Huda, M., et al.: From digital ethics to digital partnership skills: driving a safety strategy to expand the digital community?. In: Digital Transformation for Business and Society, pp. 292–310. Routledge (2023)
27. Susanto, H., et al.: Digital ecosystem security issues for organizations and governments: digital ethics and privacy. In: Web 2.0 and Cloud Technologies for Implementing Connected Government, pp. 204–228. IGI Global (2021)
28. Huda, M., Sutopo, L., Liberty, Febrianto, Mustafa, M.C.: Digital information transparency for cyber security: critical points in social media trends. In: Arai, K. (eds.) Advances in Information and Communication. FICC 2022. Lecture Notes in Networks and Systems, vol. 439, pp. 814–831. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-98015-3_55

29. Saputra, A.A., Fasa, M.I., Ambarwati, D.: Islamic-based digital ethics: the phenomenon of online consumer data security. Share. Jurnal Ekonomi dan Keuangan Islam, **11**(1), 105–128 (2022)
30. Boersma, K., Büscher, M., Fonio, C.: Crisis management, surveillance, and digital ethics in the COVID-19 era. J. Contingencies Crisis Manag. **30**(1), 2–9 (2022)
31. Fuchs, C.: Digital Ethics: Media, Communication and Society Volume Five. Taylor & Francis (2022)
32. Koivunen, S., Sahlgren, O., Ala-Luopa, S., Olsson, T.: Pitfalls and tensions in digitalizing talent acquisition: an analysis of HRM professionals' considerations related to digital ethics. Interact. Comput. iwad018 (2023)
33. Hawamdeh, M., Altınay, Z., Altınay, F., Arnavut, A., Ozansoy, K., Adamu, I.: Comparative analysis of students and faculty level of awareness and knowledge of digital citizenship practices in a distance learning environment: case study. Educ. Inf. Technol. **27**(5), 6037–6068 (2022)
34. Adeliant, J.P., Wibowo, T.O., Febrina, A.F., Roxanne, C., Kartadibrata, G.C., Syafuddin, K.: The role of communication technology in building social interaction and increasing digital literacy in Ibu Sibuk community. Jurnal Multidisiplin Madani **3**(8), 1704–1711 (2023)
35. Aziz, N.A.A., Razali, F.M., Saari, C.Z.: Penggunaaan Media Sosial dari Perspektif Psiko Spiritual Islam. Firdaus Journal **2**(1), 65–75 (2022)
36. Zainul, Z., Rasyid, A., Nasrul, W.: Distance learning model for Islamic religious education subjects in non internet server provider (ISP) areas. Firdaus J. **1**(1), 45–53 (2021)
37. Rahman, M.F.F.A., Shafie, S., Mat, N.A.A.: Penerapan Konsep QEI Dan Teknologi dalam Pembangunan Kit Letti bagi Topik Integer Matematik Ting-katan 1. Firdaus J. **2**(2), 13–25 (2022)
38. Zvereva, E.: Digital ethics in higher education: modernizing moral values for effective communication in cyberspace. Online J. Commun. Media Technol. **13**(2), e202319 (2023)
39. Huda, M.: Trust as a key element for quality communication and information management: insights into developing safe cyber-organisational sustainability. Int. J. Organ. Anal. (2023). https://doi.org/10.1108/IJOA-12-2022-3532
40. Syofiarti, S., Saputra, R., Lahmi, A., Rahmi, R.: The use of audiovisual media in learning and its impact on learning outcomes of Islamic cultural history at Madrasah Tsanawiyah Negeri 4 Pasaman. Firdaus J. **1**(1), 36–44 (2021). https://doi.org/10.37134/firdaus.vol1.1.4.2021
41. Huda, M.: Towards digital access during pandemic age: better learning service or adaptation struggling? Foresight **25**(1), 82–107 (2023). https://doi.org/10.1108/FS-09-2021-0184
42. Huda, M.: Digital marketplace for tourism resilience in the pandemic age: voices from budget hotel customers. Int. J. Organ. Anal. **31**(1), 149–167 (2023). https://doi.org/10.1108/IJOA-10-2021-2987
43. Zainul, Z., Rasyid, A., Nasrul, W.: Distance learning model for Islamic religious education subjects in non Internet Server Provider (ISP) areas. Firdaus J. **1**(1), 45–53 (2021). https://doi.org/10.37134/firdaus.vol1.1.5.2021
44. Huda, M.: Between accessibility and adaptability of digital platform: investigating learners' perspectives on digital learning infrastructure. High. Educ. Ski. Work. Based Learn. (2023). https://doi.org/10.1108/HESWBL-03-2022-0069
45. Huda, M., et al.: Strategic role of trust in digital communication: critical insights into building organizational sustainability. In: Arai, K. (ed.) FTC 2023. LNNS, vol. 3, pp. 387–403. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-47457-6_25
46. Abdul Aziz, N.A., Mohd Razali, F., Saari, C.Z.: Penggunaaan Media Sosial dari Perspektif Psiko Spiritual Islam. Firdaus J. **2**(1), 65–75 (2022). https://doi.org/10.37134/firdaus.vol2.1.6.2022

47. Hanafi, H.F., et al.: Review of learner's model for programming in teaching and learning. J. Adv. Res. Appl. Sci. Eng. Technol. **33**(3), 169–184 (2024)
48. Huda, M., Bakar, A.: Culturally responsive and communicative teaching for multicultural integration: qualitative analysis from public secondary school. Qual. Res. J. (2024). https://doi.org/10.1108/QRJ-07-2023-0123
49. Syafri, N., Ali, A.H., Ramli, S.: Penggunaan kaedah Inovasi Sambung dan Baca Bahasa Arab (SaBBAr) dalam meningkatkan kemahiran membaca perkataan Bahasa Arab murid di Sekolah Rendah. Firdaus J. **2**(2), 62–71 (2022). https://doi.org/10.37134/firdaus.vol2.2.6.2022
50. Huda, M., et al.: Enhancing digital leadership direction: insight into empowering gender violence prevention. In: Mishra, D., Ngoc Le, A., McDowell, Z. (eds.) Communication Technology and Gender Violence. Signals and Communication Technology, pp. 147–164. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-45237-6_13
51. Ab Rahim, N.M.Z., Saari, Z., Mohamad, A.M., Rashid, M.H., Mohamad Norzilan, N.I.: Konsep Ulul Albab dalam Al-Quran dan hubungannya dengan pembelajaran kursus Sains, Teknologi dan Manusia di UTM Kuala Lumpur. Firdaus J. **2**(2), 72–78 (2022). https://doi.org/10.37134/firdaus.vol2.2.7.2022
52. Huda, M., et al.: Understanding of digital ethics for information trust: a critical insight into gender violence anticipation. In: Mishra, D., Ngoc Le, A., McDowell, Z. (eds.) Communication Technology and Gender Violence. Signals and Communication Technology, pp. 165–181. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-45237-6_14
53. Hanafi, H.F., Wahab, M.H.A., Selamat, A.Z., Masnan, A.H., Huda, M.: A systematic review of augmented reality in multimedia learning outcomes in education. In: Singh, M., Kang, D.K., Lee, J.H., Tiwary, U.S., Singh, D., Chung, W.Y. (eds.) IHCI 2020. LNCS, vol. 12616, pp. 63–72. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-68452-5_7
54. Khalili, Huda, M., Rosman, A.S., Mohamed, A.K., Marni, N.: Digital learning enhancement through social network site (SNS). In: Silhavy, R., Silhavy, P., Prokopova, Z. (eds.) CoMeSySo 2021. LNNS, vol. 232, pp. 421–431. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-90318-3_35
55. Leh, F.C., Anduroh, A., Huda, M.: Level of knowledge, skills and attitude of trainee teachers on Web 2.0 applications in teaching geography in Malaysia schools. Heliyon **7**(12) (2021)
56. Sukadari, S., Huda, M.: Culture sustainability through co-curricular learning program: learning batik cross review. Educ. Sci. **11**(11), 736 (2021)
57. Huda, M., Gusmian, I., Mulyo, M.T.: Towards eco-friendly responsibilities: Indonesia field school model cross review. J. Comp. Asian Dev. (JCAD) **18**(2), 1–12 (2021)
58. Rachman, A., Oktoviani, I., Manurung, P.: Digitization of zakat and charity BAZNAS Tangerang city through crowdfunding platform tangerangsedekah. id. Firdaus J. **3**(1), 96–106 (2023). https://doi.org/10.37134/firdaus.vol3.1.9.2023
59. Huda, M., et al.: Understanding modern learning environment (MLE) in big data era. Int. J. Emerg. Technol. Learn. **13**(5), 71–85 (2018). https://doi.org/10.3991/ijet.v13i05.8042
60. Mohamad Shokri, S.S., Salihan, S.: Modul pembangunan Program Huffaz Profesional Universiti Tenaga Nasional: Satu pemerhatian kepada konstruk pembangunan Al-Quran. Firdaus J. **3**(2), 12–23 (2023). https://doi.org/10.37134/firdaus.vol3.2.2.2023
61. Berlian, Z., Huda, M.: Reflecting culturally responsive and communicative teaching (CRCT) through partnership commitment. Educ. Sci. **12**(5), 295 (2022)
62. Yousefi, S., Tosarkani, B.M.: Exploring the role of blockchain technology in improving sustainable supply chain performance: a system-analysis-based approach. IEEE Trans. Eng. Manag. (2023)
63. Omar, M.N.: Inovasi pengajaran & pemudahcaraan menggunakan Aplikasi Ezi-Maq (MAHARAT AL-QURAN) untuk menarik minat pelajar menguasai Ilmu Tajwid. Firdaus J. **2**(2), 79–87 (2022). https://doi.org/10.37134/firdaus.vol2.2.8.2022

64. Huda, M., Sutopo, L., Liberty, Febrianto, Mustafa, M.C.: Digital information transparency for cyber security: critical points in social media trends. In: Arai, K. (ed.) FICC 2022. LNNS, vol. 439, pp. 814–831. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-98015-3_55
65. Borham, A.H., et al.: Information and communication ethics in social media for indigenous people's religious understanding: a critical review. In: Proceedings of World Conference on Information Systems for Business Management, ISBM 2023. Springer (2024)
66. Ali, A.H.: Pelajar Khalifah Profesional Tempaan Ulul Albab sorotan penerapannya berasaskan komponen QEI di UPSI. Firdaus J. **3**(2), 51–63 (2023). https://doi.org/10.37134/firdaus.vol3.2.5.2023
67. Jusoh, A., Huda, M., Abdullah, R., Lee, N.: Development of digital heritage for archaeo-visit tourism resilience: evidences from E-Lenggong web portal. In: Proceedings of World Conference on Information Systems for Business Management, ISBM 2023. Springer (2024)
68. Rahim, M.M.A., Huda, M., Borham, A.H., Kasim, A.Y., Othman, M.S.: Managing information leadership for learning performance: an empirical study among public school educators. In: Proceedings of World Conference on Information Systems for Business Management, ISBM 2023. Springer (2024)
69. Hehsan, A., et al.: Digital muhadathah: framework model development for digital Arabic language learning. In: Cyber Security and Applications: Proceedings of ICTCS 2023, vol. 3. Springer, Cham (2024)
70. Muharom, F., Farhan, M., Athoillah, S., Rozihan, R., Muflihin, A., Huda, M.: Digital technology skills for professional development: insights into quality instruction performance. In: ICT: Applications and Social Interfaces: Proceedings of ICTCS 2023, vol. 1. Springer, Cham (2024)
71. Muhamad, N., Huda, M., Hashim, A., Tabrani, Z.A., Maarif, M.A.: Managing technology integration for teaching strategy: public school educators' beliefs and practices. In: ICT: Applications and Social Interfaces: Proceedings of ICTCS 2023, vol. 1. Springer, Cham (2024)
72. Wahid, A., et al.: Digital technology for indigenous people's knowledge acquisition process: insights from empirical literature analysis. In: Intelligent Strategies for ICT: Proceedings of ICTCS 2023, vol. 2. Springer, Cham (2024)
73. Huda, M., Taisin, J.N., Muhamad, M., Kiting, R., Yusuf, R.: Digital technology adoption for instruction aids: insight into teaching material content. In: Intelligent Strategies for ICT: Proceedings of ICTCS 2023, vol. 2. Springer, Cham (2024)
74. Alwi, A.S.Q., Ibrahim, R.: Isu terhadap Penggunaan Teknologi Media Digital dalam kalangan guru pelatih jurusan Pendidikan Khas. Firdaus J. **2**(2), 88–93 (2022). https://doi.org/10.37134/firdaus.vol2.2.9.2022
75. Zamri, F.A., Muhamad, N., Huda, M., Hashim, A.: Social media adoption for digital learning innovation: insights into building learning support. In: Cyber Security and Applications: Proceedings of ICTCS 2023, vol. 3. Springer, Cham (2024)
76. Huda, M., et al.: Trust in electronic record management system: insights from Islamic-based professional and moral engagement-based digital archive. In: Software Engineering Methods in Systems and Network Systems - Proceedings of 7th Computational Methods in Systems and Software 2023. Springer, Cham (2024)
77. Yahya, S.F., Othman, M.A.: Penggunaan video dalam Pengajaran dan Pembelajaran Pendidikan Moral Tingkatan 2. Firdaus J. **2**(2), 94–105 (2022). https://doi.org/10.37134/firdaus.vol2.2.10.2022
78. Rahman, M.H.A., Jaafar, J., Huda, M.: Information and communication skills for higher learners competence model. In: Software Engineering Methods in Systems and Network Systems - Proceedings of 7th Computational Methods in Systems and Software 2023. Springer, Cham (2024)

79. Cita Sari, D., et al.: Transformation of artificial intelligence in Islamic edu with Ulul Albab Value (global challenge perspective). Firdaus J. **3**(1), 1–9 (2023). https://doi.org/10.37134/firdaus.vol3.1.1.2023
80. Huda, M., et al.: Big data emerging technology: insights into innovative environment for online learning resources. Int. J. Emerg. Technol. Learn. **13**(1), 23–36 (2018). https://doi.org/10.3991/ijet.v13i01.6990
81. Zamri, F.A., Muhamad, N., Huda, M.: Information and communication technology skills for instruction performance: beliefs and experiences from public school educators. In: Software Engineering Methods in Systems and Network Systems - Proceedings of 7th Computational Methods in Systems and Software 2023. Springer, Cham (2024)
82. Masud, A., Borham, A.H., Huda, M., Rahim, M.M.A., Husain, H.: Managing information quality for learning instruction: insights from public administration officers' experiences and practices. In: Software Engineering Methods in Systems and Network Systems - Proceedings of 7th Computational Methods in Systems and Software 2023. Springer, Cham (2024)
83. Huda, M., et al.: Digital record management in Islamic education institution: current trends on enhancing process and effectiveness through learning technology. In: Software Engineering Methods in Systems and Network Systems - Proceedings of 7th Computational Methods in Systems and Software 2023. Springer, Cham (2024)
84. Mohd Nawi, M.Z.: Media variations in education in Malaysia: a 21st century paradigm. Firdaus J. **3**(1), 77–95 (2023). https://doi.org/10.37134/firdaus.vol3.1.8.2023
85. Huda, M., et al.: From technology adaptation to technology adoption: an insight into public school administrative management. In: Proceedings of Ninth International Congress on Information and Communication Technology, ICICT 2024. Lecture Notes in Networks and Systems. Springer, Singapore (2024)
86. Tan, A.A., Huda, M., Rohim, M.A., Hassan, T.R.R., Ismail, A.: Chat GPT in supporting education instruction sector: an empirical literature review. In: Proceedings of Ninth International Congress on Information and Communication Technology, ICICT 2024. Lecture Notes in Networks and Systems. Springer, Singapore (2024)
87. Wahid, A., et al.: Augmented reality model in supporting instruction process: a critical review. In: Proceedings of Ninth International Congress on Information and Communication Technology, ICICT 2024. Lecture Notes in Networks and Systems. Springer, Singapore (2024)
88. Musolin, M.H., Serour, R.O.H., Hamid, S.A., Ismail, A., Huda, M., Rohim, M.A.: Developing personalized Islamic learning in digital age: pedagogical and technological integration for open learning resources (OLR). In: Proceedings of Ninth International Congress on Information and Communication Technology, ICICT 2024. Lecture Notes in Networks and Systems. Springer, Singapore (2024)
89. Musolin, M.H., Ismail, M.H., Huda, M., Hassan, T.R.R., Ismail, A.: Towards an Islamic education administration system: a critical contribution from technology adoption. In: Proceedings of Ninth International Congress on Information and Communication Technology, ICICT 2024. Lecture Notes in Networks and Systems. Springer, Singapore (2024)
90. Musolin, M.H., Ismail, M.H., Farhan, M., Rois, N., Huda, M., Rohim, M.A.: Understanding of artificial intelligence for Islamic education support and service: insights from empirical literature review. In: Proceedings of Ninth International Congress on Information and Communication Technology, ICICT 2024. Lecture Notes in Networks and Systems. Springer, Singapore (2024)
91. Susilowati, T., et al.: Getting parents involved in child's school: using attendance application system based on SMS gateway. Int. J. Eng. Technol. **7**(2.27), 167–174 (2018)
92. Wahyudin, U., Jandra, M., Huda, M., Maseleno, A.: Examining development quality practice in higher education: evidence from islamic higher education institution (IHEI) in Indonesia. Test Eng. Manag. **81**(11–12), 4298–4310 (2019)

93. Susilowati, T., et al.: Learning application of Lampung language based on multimedia software. Int. J. Eng. Technol. **7**(2.27), 175–181 (2018)
94. Abadi, S., et al.: Application model of k-means clustering: insights into promotion strategy of vocational high school. Int. J. Eng. Technol. **7**(2.27), 182–187 (2018)
95. Huda, M., et al.: Towards digital servant leadership for organisational stability: driving processes in the pandemic age? In: Digital Transformation for Business and Society: Contemporary Issues and Applications in Asia. Taylor & Francis, UK (2023b)
96. Aminudin, N., et al.: The family hope program using AHP method. Int. J. Eng. Technol. **7**(2.27), 188–193 (2018)
97. Wulandari, et al.: Design of library application system. Int. J. Eng. Technol. **7**(2.27), 199–204 (2018)
98. Aminudin, N., et al.: Higher education selection using simple additive weighting. Int. J. Eng. Technol. **7**(2.27), 211–217 (2018)
99. Zainuri, A., Huda, M.: Empowering cooperative teamwork for community service sustainability: insights from service learning. Sustainability **15**(5), 4551 (2023)
100. Maseleno, A., et al.: Hau-Kashyap approach for student's level of expertise. Egypt. Inform. J. **20**(1), 27–32 (2019)
101. Latif, M.K., Md Saad, R., Abd Hamid, S.: Islamic education teachers' competency in teaching Qiraat Sab'ah for the Quranic Class. Firdaus J. **3**(1), 19–27 (2023). https://doi.org/10.37134/firdaus.vol3.1.3.2023
102. Huda, M., et al.: Empowering Islamic-based digital competence and skills: how to drive it into reconstructing safety strategy from gender violence. In: Software Engineering Methods in Systems and Network Systems - Proceedings of 7th Computational Methods in Systems and Software 2023. Springer, Cham (2024)
103. Richey, R., Klein, J.: Design and Development Research: Methods, Strategies and Issues. Lawrence Erlbaum Associates, Mahwah (2007)
104. Huda, M., Kartanegara, M.: Islamic spiritual character values of al-Zarnūjī's Ta 'līm al-Muta 'allim. Mediterranean J. Soc. Sci. **6**(4S2), 229–235 (2015)
105. Huda, M., Yusuf, J.B., Jasmi, K.A., Nasir, G.A.: Understanding comprehensive learning requirements in the light of al-Zarnūjī's Ta'līm al-Muta'allim. SAGE Open **6**(4), 1–14 (2016)
106. Huda, M., Yusuf, J.B., Jasmi, K.A., Zakaria, G.N.: Al-Zarnūjī's concept of knowledge ('ilm). SAGE Open **6**(3), 1–13 (2016)
107. Huda, M., Jasmi, K.A., Mohamed, A.K., Wan Embong, W.H., Safar, J.: Philosophical investigation of Al-Zarnuji's Ta'lim al-Muta'allim: strengthening ethical engagement into teaching and learning. Soc. Sci. **11**(22), 5516–551 (2016c)
108. Huda, M., et al.: Innovative teaching in higher education: the big data approach. Turk. Online J. Educ. Technol. **15**(Spec. Issue), 1210–1216 (2016)
109. Othman, R., Shahrill, M., Mundia, L., Tan, A., Huda, M.: Investigating the relationship between the student's ability and learning preferences: evidence from year 7 mathematics students. New Educ. Rev. **44**(2), 125–138 (2016)
110. Anshari, M., Almunawar, M.N., Shahrill, M., Wicaksono, D.K., Huda, M.: Smartphones usage in the classrooms: learning aid or interference? Educ. Inf. Technol. **22**(6), 3063–3079 (2017)
111. Huda, M., Sabani, N., Shahrill, M., Jasmi, K.A., Basiron, B., Mustari, M.I.: Empowering learning culture as student identity construction in higher education. In: Shahriar, A., Syed. G. (eds.) Student Culture and Identity in Higher Education, pp. 160–179. IGI Global, Hershey (2017a). https://doi.org/10.4018/978-1-5225-2551-6.ch010
112. Huda, M., et al.: Empowering children with adaptive technology skills: careful engagement in the digital information age. Int. Electron. J. Elem.Tary Educ. **9**(3), 693–708 (2017)

113. Huda, M., Shahrill, M., Maseleno, A., Jasmi, K.A., Mustari, I., Basiron, B.: Exploring adaptive teaching competencies in big data era. Int. J. Emerg. Technol. Learn. **12**(3), 68–83 (2017c)
114. Huda, M., Jasmi, K.A., Basiran, B., Mustari, M.I.B., Sabani, A.N.: Traditional wisdom on sustainable learning: an insightful view from Al-Zarnuji's Ta 'lim al-Muta 'allim. SAGE Open **7**(1), 1–8 (2017)
115. Huda, M., Jasmi, K.A., Alas, Y., Qodriah, S.L., Dacholfany, M.I., Jamsari, E.A.: Empowering civic responsibility: insights from service learning. In: Burton, S. (ed.) Engaged Scholarship and Civic Responsibility in Higher Education, pp. 144–165. IGI Global, Hershey (2017e). https://doi.org/10.4018/978-1-5225-3649-9.ch007
116. Huda, M., et al.: Innovative e-therapy service in higher education: mobile application design. Int. J. Interact. Mob. Technol. **11**(4), 83–94 (2017g)
117. Huda, M., Jasmi, K.A., Alas, Y., Qodriah, S.L., Dacholfany, M.I., Jamsari, E.A.: Empowering civic responsibility: isights from service learning. In: Burton, S. (ed.) Engaged Scholarship and Civic Responsibility in Higher Education, pp. 144–165. IGI Global, Hershey (2017h). https://doi.org/10.4018/978-1-5225-3649-9.ch007
118. Aminin, S., Huda, M., Ninsiana, W., Dacholfany, M.I.: Sustaining civic-based moral values: insights from language learning and literature. Int. J. Civ. Eng. Technol. **9**(4), 157–174 (2018)
119. Huda, M., Teh, K.S.M.: Empowering professional and ethical competence on reflective teaching practice in digital era. In: Dikilitas, K., Mede, E., Atay D. (eds.) Mentorship Strategies in Teacher Education, pp. 136–152. IGI Global, Hershey (2018). https://doi.org/10.4018/978-1-5225-4050-2.ch007
120. Huda, M., Teh, K.S.M., Nor, N.H.M., Nor, M.B.M.: Transmitting leadership based civic responsibility: insights from service learning. Int. J. Ethics Syst. **34**(1), 20–31 (2018). https://doi.org/10.1108/IJOES-05-2017-0079
121. Huda, M., Mulyadi, D., Hananto, A.L., Nor Muhamad, N.H., Mat Teh, K.S., Don, A.G.: Empowering corporate social responsibility (CSR): insights from service learning. Soc. Responsib. J. **14**(4), 875–894 (2018)
122. Maseleno, A., et al.: Mathematical theory of evidence to subject expertise diagnostic. ICIC Express Lett. **12**(4), 369 (2018a). https://doi.org/10.24507/icicel.12.04.369
123. Maseleno, A., Sabani, N., Huda, M., Ahmad, R., Jasmi, K.A., Basiron, B.: Demystifying learning analytics in personalised learning. Int. J. Eng. Technol. **7**(3), 1124–1129 (2018)
124. Huda, M., Sudrajat, S., Kawangit, R.M., Teh, K.S.M., Jalal, B.: Strengthening divine values for self-regulation in religiosity: insights from Tawakkul (trust in God). Int. J. Ethics Syst. **35**(3), 323–344 (2019). https://doi.org/10.1108/IJOES-02-2018-00257
125. Fitrian, Y., et al.: Application design for determining suitable cosmetics with the facial skin type using fuzzy logic approach. J. Comput. Theor. Nanosci. **16**(5–6), 2153–2158 (2019)
126. Sukadari, Jandra, M., Sutarto, Hehsan, A., Junaidi, J., Huda, M.: Exploring specific learning difficulties in primary schools: an empirical research. Test Eng. Manag. **81**(11–12), 4387–4399 (2019)
127. Susilowati, T., et al.: Decision support system for determining lecturer scholarships for doctoral study using CBR (case-based reasoning) method. Int. J. Recent. Technol. Eng. **8**(1), 3281–32 (2019)
128. Muslihudin, M., et al.: Decision support system in kindergarten selection using TOPSIS method. Int. J. Recent. Technol. Eng. **8**(1), 3291–3298 (2019)
129. Abadi, S., et al.: Identification of sundep, leafhopper and fungus of paddy by using fuzzy SAW method. Int. J. Pharm. Res. **11**(1), 695–699 (2019)
130. Maseleno, A., Shankar, K., Huda, M., Othman, M., Khoir, P., Muslihudin, M.: Citizen Economic Level (CEL) using SAW. In: Expert Systems in Finance: Smart Financial Applications in Big Data Environments, vol. 97. Routledge, New York (2019)

131. Huda, M., et al.: Empowering technology use to promote virtual violence prevention in higher education context. In: Intimacy and Developing Personal Relationships in the Virtual World, pp. 272–291. IGI Global, Hershey (2019). https://doi.org/10.4018/978-1-5225-4047-2.ch015
132. Kencana, U., Huda, M., Maseleno, A.: Waqf administration in historical perspective: evidence from Indonesia. Test Eng. Manag. **81**(11–12) (2019)
133. Huda, M., et al.: Learning quality innovation through integration of pedagogical skill and adaptive technology. Int. J. Innov. Technol. Explor. Eng. **8**(9S3), 1538–1541 (2019)
134. Salamah, P., Jandra, M., Sentono, T., Huda, M., Maseleno, A.: The effects of emotional intelligence, family environment and learning styles on social-science learning outcomes: an empirical analysis. Test Eng. Manag. **81**(11–12), 4374–4386 (2019)
135. Tarto, J.J., Huda, M., Maseleno, A.: Expanding trilogy-based headmaster leadership: a conceptual framework. Test Eng. Manag. **81**(11–12), 4356–4373 (2019)

Cyber Security Management in Metaverse: A Review and Analysis

Farnaz Farid[1](✉), Abubakar Bello[1], Nusrat Jahan[2], and Razia Sultana[2]

[1] School of Social Sciences, Western Sydney University, Sydney, Australia
farnaz.farid@westernsydney.edu.au
[2] Department of Computer Science and Engineering, Eastern University, Chennai, India

Abstract. Metaverse is an instance of the future web. It describes a collective virtual shared space, where users can engage with virtual environments in real-time. Using Artificial Intelligence (AI), Virtual Reality (VR), Augmented Reality (AR), digital twin, and blockchain technology, the Metaverse is becoming a reality in science fiction. With these technologies, the Metaverse will inevitably emerge with new cybersecurity challenges and risks. Thus, significant privacy and security breaches in the Metaverse could hamper its widespread deployment. This review and analysis paper attempts to dissect the cybersecurity threats related to the Metaverse by outlining the supporting technologies and exploring the cybersecurity vulnerabilities in each of these technologies. The motive is to assess security challenges, privacy threats, and potential controls to manage virtual reality environments with real-time ambience.

Keywords: Metaverse · VR · AR · Digital Twin · Blockchain · Cybersecurity in Metaverse · Security management in Metaverse · Metaverse data privacy

1 Introduction

In this era, some advanced technologies have brought about significant impacts by enabling new and improved modes of communication, collaboration, and commerce. Immersive technologies such as Virtual Reality (VR), Augmented Reality (AR), and extended reality (XR) [1, 2] are now being developed with the goal of creating the next ubiquitous computing paradigm, which could revolutionize business, education, distance work, and entertainment. The Metaverse, presented as a fresh take on the Internet, is quickly becoming a promising commercial frontier for various industries [3–5]. The Metaverse has the potential to bring together people from all over the world, regardless of their physical location, from socializing and gaming to learning and business. It could also facilitate new forms of creativity, such as virtual art and music, and new business models, such as virtual real estate and digital fashion.

All technology confederates, such as information systems and virtual environment designers, engineers and third-party app developers, are required to introduce inventive security systems for the Metaverse. This includes building the Metaverse itself. Apart from this, the Metaverse carries the risk of exacerbating the security vulnerabilities

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 183–193, 2024.
https://doi.org/10.1007/978-3-031-53552-9_16

already present in all of the technologies that make it possible, including social networks, cloud computing, AR, VR, and XR [6].

Furthermore, some security issues, such as social engineering attacks, ransomware, keylogging attacks [7], and network credential theft, may occur because of the AR and VR technologies that enable the Metaverse. A user's identity within the Metaverse may be compromised if programmers abuse security gaps in these virtual environments [8]. This highlights the growing significance of cyber security measures. This paper focuses on examining and validating the existence of the Metaverse's security and privacy flaws and gaps.

Section 2 of the paper provides an overview of the Metaverse, Sect. 3 discusses some key technologies used in the Metaverse, Sect. 4 describes the relationship between the Metaverse and VR, Sects. 5 and 6 depict the security challenges of the Metaverse and cyber security controls for the Metaverse, Sect. 7 discusses the overall research question in light of the review and analysis of the Metaverse and provides with some recommendations. Finally, Sect. 8 concludes the paper.

2 Evolution of the Metaverse

The word "metaverse" (Meta: Greek for "after" or "beyond" and universe) was popularised by Neal Stephenson in his 1992 science fiction novels and films like "Snow Crash" and "Ready Player One," which take place in a virtual environment where humans, represented by programmed avatars, interact with one other and software agents in a third-person perspective. Essentially, the Metaverse is a persistent multiuser environment that blurs the lines between the real and virtual worlds [9, 10]. It may be a speculative advancement of the web in which all perspectives of online life are coordinated into a single virtual reality space. The Metaverse could be observed by the following cycle of the web: a uniquely combined, interactive, committed, virtual 3D environment [11] shared by multiple individuals, life-logging [12] where humans experience life in ways they could not live in the physical world.

The first and second industrial revolutions instituted significant innovation in the physical world, but with the emergence of powerful technologies in the third and fourth industrial revolutions, the virtual world is undergoing a wave of innovation [13]. Many major internet companies, including Facebook, Microsoft, Tencent, and NVIDIA, have lately indicated their intentions to enter the Metaverse. In particular, Facebook focuses on creating the future Metaverse by rebranding itself as "Meta" [14]. Also, the ability to manipulate, exchange, and update 3D twins of advanced machines exported to Mixed Reality (MR) creates huge benefits for aviation institutions [15].

3 Metaverse and Emerging Technologies

Some immersive technologies encompass presenting the client with a visual representation of virtual substance within the Metaverse. The development of the Metaverse relies on a combination of hardware, software, and networking technologies to create immersive virtual experiences that can be accessed and interacted with by users from all over the

world [16, 17]. AR, VR, haptic and brain-computer interfaces, intelligent sensors, cryptocurrency, blockchain, adequate bandwidth or interoperability standards, holograms, and digital twins are among the fundamental technologies necessary for the successful launch of the Metaverse [18]. Figure 1 illustrates the technologies in Metaverse.

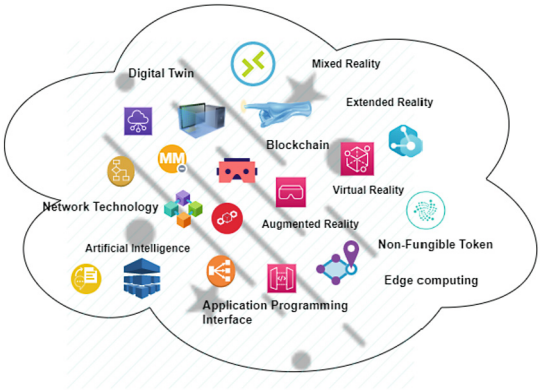


Fig. 1. Technologies in Metaverse

Multimodal Interactive Technology: XR broadly utilizes VR, AR, and MR innovation to supply multi-sensory inundation, increased encounters, and real-time client, avatar, and environment interaction. Head Mounted Displays (HMDs), wearable gadgets, XR gloves, and brain-computer interaction (BCI), control advanced avatars within the Metaverse [14, 19].

Artificial Intelligence: For object recognition, image processing, speech recognition and processing, and recognizing human activities, Artificial Intelligence (AI), Machine Learning (ML), Deep Learning (DL), Natural Language Processing (NLP), and Computer Vision are utilized in the Metaverse [20].

Digital Twin: In the Metaverse, a digital twin can reflect the physical world into the virtual one [21]. It empowers the reflection of physical substances and the expectation of their virtual bodies by inspecting live streams of tangible information, physical models, and chronicled data.

Network Technology: Some network technologies, such as 6G, 5G, Software-defined networking (SDN), and space-air-ground integrated network (SAGIN) [22], are immersive networking and communication technologies in the Metaverse. The utilization of 5G and 6G networks assures dependable and efficient communication for the vast array of devices in the Metaverse while enhancing overall mobility assistance. SDN allows flexibility, scalability, and dynamic network administration to run a vast metaverse web.

Application Programming Interface (API): Oculus, and Facebook's VR division have released the Passthrough API. Various interoperable APIs are critical for ensuring that users can use avatars and digital assets across metaverses, allowing for teleportation between various virtual worlds.

Blockchain: This innovation works as a stockroom, which permits clients to stock information in any place within the Metaverse. Blockchain innovation has the potential

to offer a total financial framework that interfaces the Metaverse's virtual world and the genuine world [23].

Non-Fungible Token (NFT): It is a type of cryptocurrency produced by Ethereum smart contracts, enabling the use of novel virtual economy transactions and systems. These NFTs, as well as cryptocurrencies, would have widespread applications and also a medium of value interchange in the Metaverse [14].

Edge computing: alludes to the empowering innovations that permit computation to be performed at the edge of the arranged, primarily downstream information in cloud administrations and upstream information for IoT administrations. By facilitating individual information on rented cloud capacity, edge computing permits this to be put away on individual gadgets, consequently lessening information spills and improving cyber posture.

4 The Relationship Between Metaverse and Virtual Reality

Meta is undertaking the establishment of a social network for VR, while Roblox is enabling user-generated video games. [25]. The combination of social media platforms, online gaming, AR, VR, XR, and cryptocurrencies is known as the Metaverse [26]. The user experience can be improved by applying augmented reality by superimposing visual components, sound, and other information onto real-world situations. In another dimension, virtual reality comprises virtual environments and augments fictional realities [27].

Metaverse refers to a simulated, immersive world, which was made possible by developing technologies such as virtual and augmented reality. Users create their unique graphical representations and avatars, which they then use to navigate the many different virtual worlds available within the Metaverse. The way to explore these virtual worlds is virtual reality (VR). Individuals can encounter what they are pursuing in genuine time by utilizing VR hardware such as VR headsets and gloves. The sensors can take after the user's developments, and at the same time, telepresence is built up. Increased reality acts as a bridge between the virtual world and the genuine world. However, virtual reality empowers us to visit three-dimensional areas [13, 28].

5 Metaverse Security Challenges

The Metaverse, which is immersed as a popular and promising technology in the digital age, has unique characteristics that differentiate it from other tools, making it crucial in various fields like education, business, and digital medicine. However, its use may also lead to an increase in cybersecurity vulnerabilities and risks. With millions of cyberattacks propelled daily, securing information within the Metaverse remains challenging. The utilization of AR and VR advances within the Metaverse may grant rise to security concerns ranging from social engineering attacks, ransomware attacks, organized credential theft, and identity/character theft. Exploiting these security vulnerabilities in these gadgets, cybercriminals may be able to compromise or hijack a user's personality within the Metaverse [27].

Despite the Metaverse's promising advantages, protection and security concerns remain the most significant impediments blocking its progression. The risks of handling massive data flows, widespread user identification practices, inequitable outcomes produced by AI algorithms, and the security of the tangible infrastructure and humans are only a few examples of the broad spectrum of security vulnerabilities and invasions of privacy that may occur in the Metaverse. Emerging technology incidents have included the theft of virtual currencies, the appropriation of wearable technology or cloud storage, and the misuse of AI to create false news [29].

The paper [30] pointed out the risk of following client behaviour, counting social systems and savvy homes, and contends that the Metaverse may also permit the same kind of client following, raising extreme security issues. User information, communications, situations, and items were the four main categories utilized to assess security concerns in a subsequent analysis of the security of the Metaverse. There were also some recommended defences [31]. As clients within the Metaverse must be interestingly recognized, headsets, VR glasses, and other devices may be unlawfully utilized to track users' real-world areas [32]. Last but not least, hackers could use corrupted gadgets as access points into real gear like domestic appliances to endanger personal safety. They can also use advanced continuous risk attacks to threaten critical systems like grid power structures, elevated rail networks, and water distribution networks [14].

To identify known vulnerabilities and assess the cyber security and data privacy risks related to the metaverse system, it is essential to comprehend all of those technologies and their security flaws:

1. ***Identification and authentication:*** Users can build several digital identities in a metaverse to utilize on various virtual platforms. As a result, it may be challenging to confirm users' identities. Also, managing and securely keeping the authentication credentials connected to these identities may provide difficulties.
2. ***Data Privacy:*** In the Metaverse, data privacy poses the greatest threat to cyber security. There are privacy issues with the user data (also critical biometric data) that is gathered and the uses that are made of it. The user data acquired would be subject to data privacy directives depending on the end users' location.
3. ***Access Risk:*** Users can choose which data will be shared based on the privacy directives. Businesses must ensure that information about user data collection, storage, and purposes is disclosed. Also, it is required to put in restrictions for goal, content, and collection limitation, in addition to confidentiality and integrity.
4. ***Security of NFTs and Blockchain:*** There have been instances of vulnerabilities in Blockchain smart contract platforms that malicious actors have exploited. Controls such as verifying the caller's identity through methods like signature verification are suggested to mitigate vulnerabilities.
5. ***Platform/Application code vulnerabilities:*** To ensure software development security, it is essential to follow industry standards such as OWASP (Open Web Application Security Project) and BSI (British Standards Institution) and conduct regular code reviews.
6. ***Algorithmic fairness:*** As artificial intelligence is more widely used, algorithms play a bigger part in actual decision-making in modern life, including judgments about credit and insurance. Knowing the factors that influence algorithmic decisions and

then calculating the effect of each influencing input factor, including gender or racial background, on the system's ultimate choice [33].

7. ***API/Sensor security:*** APIs are a target for attackers because they expose application logic and sensitive data, including Personally Identifiable Information.
8. ***Data Center/Cloud Security:*** Some of the primary measures in the cloud consist of encryption of cloud data, regulation of access, evaluation of security risks, authentication and authorization procedures, safeguarding of media, contingency strategies, protection of privacy, adherence to legal and regulatory standards, management of data centre operations, execution of incident response plans, improved awareness, and provision of training.
9. ***Distributed Denial of Service:*** DDoS attacks also occur in the Metaverse. Technological measures include Content Delivery Networks (CDNs) for DDoS Protection, Bot Mitigation, Web Application Firewalls (WAFs), and API Gateways for Network Layer [17].
10. ***Verification and Credibility:*** A pragmatic instance of the issues brought on by computations and robotization is the security challenges of content integrity and client validation. Human-machine communications will become more common in the Metaverse, assuming they are not required for some professions.
11. ***Radicalization and Polarization in Singleton's Universe:*** The singleton nature of the Metaverse will lead to security issues. A metaverse is a vast collection of users, objects, services, and applications. Its ability to serve as a focal route for such content does not guarantee its prosperity. A small group of enormous metaverses will replace most current Web stages.
12. ***Gadget Insecurity:*** By using the web, programmers can intercept data travelling through head-mounted devices connected to the Metaverse. Under the assumption that the attacker shares realistic or explicit content, this could be considered a social assault. In addition, if they post hate speech or other deceptive advertisements, it might be seen as a political tool [34].

6 Analysis of the Cyber Security Controls for the Metaverse

As the utilization of metaverse situations proceeds to develop, the requirements for compelling cyber security controls become progressively imperative. Centralized character, Unified character, and Self-sovereign personality (SSI) [5] are the three computerized characters for the Metaverse [14]. Customary security reviews and powerlessness evaluations can offer assistance in distinguishing and addressing potential security shortcomings; malevolent on-screen characters can abuse them. It is crucial to consider cybersecurity measures as the Metaverse spreads to secure users and their assets. These potential security measures for the Metaverse are listed below:

Encryption: All data transported within the Metaverse should be secured using robust encryption techniques to avoid interception and unwanted access.

Access control: A strong access control mechanism should be in place to ensure that only people with permission can enter the Metaverse. Role-based access control, password rules, and two-factor authentication might be examples.

Observing: The Metaverse should always be checked for suspicious activity, such as abnormal login endeavours, changes in client behaviour, and unauthorized get-to endeavours. This may be finished by utilizing security information and event management (SIEM) devices.

Powerlessness administration: Normal defenselessness appraisals and infiltration testing should be performed to identify and remediate potential security shortcomings within the Metaverse.

Information security: Solid information security controls, such as information reinforcement and recuperation methods, should be input to avoid information misfortune and guarantee information accessibility.

Occurrence reaction: A comprehensive occurrence reaction arrangement should be input to address any cyber security episodes inside the Metaverse. This arrangement ought to incorporate steps for control, examination, and remediation.

Security: Users' security ought to be secured inside the Metaverse. This incorporates collecting, as it were, the least sum of information vital, giving clients control over their information, and straightforwardly communicating how information is collected and utilized.

Instruction and preparation: Clients should be taught to secure themselves inside the Metaverse. This incorporates preparing how to recognize and maintain a strategic distance from phishing assaults, how to set solid passwords, and how to report suspicious activity.

The existing security measures may not be foolproof and lack adaptability for metaverse platforms. The unique qualities of the Metaverse, including its remarkable nature, advanced temporal and spatial dimensions [35], long-term sustainability [36], capacity for data exchange, ability to expand [14], and diverse nature, could present challenges in adequately ensuring its security.

Many of the security concerns that face metaverse users are also faced by Internet users, including spam, malware assaults, data hacking, and privacy issues. The activity and information stream within the metaverse framework were not scrambled, the analysts were able to modify metaverse sessions and the related inputs and yields of the headsets to conduct MITM (man-in-the-middle) assaults [26].

7 Discussions and Recommendations

Many difficulties are anticipated to arise as the Metaverse continues to grow. As haptic technology progresses, the vulnerabilities concerning security within the Metaverse will become more tangible and real. Keeping biometric information secure also raises the possibility of data leaks and misuse. Further study in establishing security in the Metaverse appears necessary to solve these issues and build a safer digital future. The security strategies are outlined is depicted in Table 1.

The two main platforms used to examine the Metaverse security performance are VR headsets and smartphones. To assess a system's sturdiness, various cyber security attacks can be performed. Among the most frequent assaults are spoofing, Sybil, DDoS, and eavesdropping. The research reveals that the two widely adopted AI techniques in the realm of metaverse security are neural networks (NNs) and support vector machines

Table 1. The security strategies are recommended of the current method

Device	Risk	Assault Method	Suggestion	Method	Benchmarks	Concerns
AR/VR	Personal information, User data, Biometric data	Zero-effort and mimicking attacks	Blinkey	SVM, k-NN	FRR, FAR, ERR	Detecting the beginning and conclusion of a blinking event
XR	User data	Sybil	Gini and ABC Model	Gini and ABC Model	Delay	Insufficient computational data and incomplete models are the underlying issues
Drones	Drone data	Eavesdropping	PCSF	FL	QoS, security rate, training time	The intricacy of computations
Smartphone	Biometric data, User data	FS, DFL, and FS-GAN attacks	Visual speaker authentication	DNN SA-DTH-Net	FAR, FRR	A significant challenge in lip-reading is the recognition of speech content due to the extensive lexicon
VRLEs	Learning content	Issues with network connectivity, loss of data packets, and malicious observing of online traffic or Denial of Service attacks	Attack tree	Attack tree	Data packet loss and the magnitude of potential risk assessment score	A deficiency in regulating policy changes during virtual reality learning environments (VRLE) sessions
Cyber –Physical System	Transportation system data	CWM, WMFJ, and WMMJ attacks	SAF, RISA	SAF, RISA	Secrecy capacity	The problem of protecting against eavesdropping by using smart jammers

(SVMs). SVMs are particularly lauded for their exceptional data classification accuracy, multitasking abilities, and efficiency. These techniques translate multidimensional spaces into spaces with fewer dimensions to extract fixed features. Users can always access biometrics. They can use biometric information from smartphones, eyewear, or browsers. To build human-machine interactions, this data can be transformed using machine learning techniques and artificial intelligence [37].

8 Conclusions

This paper delves into the security obstacles and scrutiny of cyber-security measures for the Metaverse, drawing on current research. It was determined that hackers could attack the Metaverse using AI and that neural network techniques can significantly increase attack detection accuracy.

In conclusion, as the Metaverse continues to grow and become more prevalent, it is crucial to have effective cyber-security controls in place to protect users from potential threats such as data breaches, identity theft, and cyber-attacks. The review of cybersecurity controls for the Metaverse reveals that several important measures can be taken to improve security, including encryption, multi-factor authentication, secure development practices, and user security awareness training. It is also important to note that new cybersecurity challenges may arise as the Metaverse evolves that will require new approaches and solutions. Therefore, ongoing monitoring and continuous improvement of cyber security controls will be necessary to ensure the safety and security of users. Overall, the development and implementation of robust cybersecurity controls are essential for the continued growth and success of the Metaverse, and all stakeholders need to work together to achieve this goal.

References

1. Chung, K.H.Y., Li, D., Adriaens, P.: Technology-enabled financing of sustainable infrastructure: a case for blockchains and decentralized oracle networks. Technol. Forecast. Soc. Chang. **187**, 122258 (2023)
2. Lee, Y., Moon, C., Ko, H., Lee, S.H., Yoo, B.: Unified representation for XR content and its rendering method. In: The 25th International Conference on 3D Web Technology, pp. 1–10 (2020)
3. Mystakidis, S.: Metaverse. Encyclopedia, **2** (1), 486–497 (2022)
4. Sanchez, J.: Second life: an interactive qualitative analysis. In: Society for Information Technology & Teacher Education International Conference, pp. 1240–1243. Association for the Advancement of Computing in Education (AACE) (2007)
5. Grider, D., Maximo, M.: The metaverse: web 3.0 virtual cloud economies. Grayscale Res. 1–19 (2021)
6. Al Arafat, A., Guo, Z., Awad, A.: Vr-spy: a side-channel attack on virtual keylogging in vr headsets. In: 2021 IEEE Virtual Reality and 3D User Interfaces (VR), pp. 564–572. IEEE (2021)
7. Meteriz-Yıldıran, Ü., Yıldıran, N. F., Awad, A., Mohaisen, D.: A Keylogging inference attack on air-tapping keyboards in virtual environments. In: 2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), pp. 765–774. IEEE (2022)
8. Nath, K.: Evolution of the Internet from web 1.0 to Metaverse: The good, the bad and the ugly (2022)
9. Kürtünlüoğlu, P., Akdik, B., Karaarslan, E.: Security of virtual reality authentication methods in Metaverse: An overview (2022). arXiv preprint arXiv:2209.06447
10. Bibri, S.E.: The metaverse as a virtual model of platform urbanism: its converging AIoT, xreality, neurotech, and nanobiotech and their applications, challenges, and risks. Smart Cities **6**(3), 1345–1384 (2023)
11. Dionisio, J.D.N., III, W.G.B., Gilbert, R.: 3D virtual worlds and the Metaverse: Current status and future possibilities. ACM Comput. Surv. (CSUR), **45**(3), 1–38 (2013)

12. Bruun, A., Stentoft, M.L.: Lifelogging in the wild: participant experiences of using lifelogging as a research tool. In: Lamas, D., Loizides, F., Nacke, L., Petrie, H., Winckler, M., Zaphiris, P. (eds.) INTERACT 2019. LNCS, vol. 11748, pp. 431–451. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-29387-1_24
13. Gupta, A., Khan, H.U., Nazir, S., Shafiq, M., Shabaz, M.: Metaverse security: issues, challenges and a viable ZTA model. Electronics **12**(2), 391 (2023)
14. Wang, Y., et al.: A survey on metaverse: fundamentals, security, and privacy. IEEE Commun. Surv. Tutorials (2022)
15. Siyaev, A., Jo, G.S.: Towards aircraft maintenance metaverse using speech interactions with virtual objects in mixed reality. Sensors **21**(6), 2066 (2021)
16. Chow, Y.W., Susilo, W., Li, Y., Li, N., Nguyen, C.: Visualization and cybersecurity in the metaverse: a survey. J. Imaging **9**(1), 11 (2022)
17. Sebastian, G.: A descriptive study on metaverse: cybersecurity risks, controls, and regulatory framework. Int. J. Secur. Privacy Pervasive Comput. (IJSPPC) **15**(1), 1–14 (2023)
18. The Metaverse: How to access the virtual world, Infomineo, Jan 25, 2023. https://infomineo.com/the-metaverse-how-to-access-the-virtual-world/#:~:text=Augmented%20Reality%20(AR)%2C%20Virtual. Accessed 10 Apr 2023
19. Koutitas, G., Smith, S., Lawrence, G.: Performance evaluation of AR/VR training technologies for EMS first responders. Virtual Reality **25**, 83–94 (2021)
20. Huynh-The, T., Pham, Q.V., Pham, X.Q., Nguyen, T.T., Han, Z., Kim, D.S.: Artificial intelligence for the metaverse: a survey. Eng. Appl. Artif. Intell. **117**, 105581 (2023)
21. Ramu, S.P., et al.: Federated learning enabled digital twins for smart cities: concepts, recent advances, and future directions. Sustain. Cities Soc. **79**, 103663 (2022)
22. Wang, Y., Su, Z., Ni, J., Zhang, N., Shen, X.: Blockchain-empowered space-air-ground integrated networks: opportunities, challenges, and solutions. IEEE Commun. Surv. Tutorials **24**(1), 160–209 (2021)
23. Jeon, H.J., Youn, H.C., Ko, S.M., Kim, T.H.: Blockchain and AI Meet in the Metaverse. Adv. Convergence Blockchain Artif. Intell. **73** (10.5772) (2022)
24. Salloum, S., et al.: Sustainability model for the continuous intention to use metaverse technology in higher education: a case study from Oman. Sustainability **15**(6), 5257 (2023)
25. Ravenscraft, E.: What is the Metaverse, exactly. Everything you never wanted to know about the future of talking about the future (2022). https://www.wired.com/story/what-is-the-metaverse
26. Qamar, S., Anwar, Z., Afzal, M.: A systematic threat analysis and defense strategies for the metaverse and extended reality systems. Comput. Secur. 103127 (2023)
27. Jaipong, P., et al.: A review of metaverse and cybersecurity in the digital era. Int. J. Comput. Sci. Res. 1125–1132 (2023)
28. Rose, S., Borchert, O., Mitchell, S., Connelly, S.: Zero trust architecture (No. NIST Special Publication (SP) 800–207). National Institute of Standards and Technology (2020)
29. Leenes, R.: Privacy in the metaverse. In: Fischer-Hübner, S., Duquenoy, P., Zuccato, A., Martucci, L. (eds.) Privacy and Identity 2007. ITIFIP, vol. 262, pp. 95–112. Springer, Boston, MA (2008). https://doi.org/10.1007/978-0-387-79026-8_7
30. Falchuk, B., Loeb, S., Neff, R.: The social metaverse: battle for privacy. IEEE Technol. Soc. Mag. **37**(2), 52–61 (2018)
31. Zhao, R., Zhang, Y., Zhu, Y., Lan, R., Hua, Z.: Metaverse: Security and privacy concerns (2022). arXiv preprint arXiv:2203.03854
32. Shang, J., Chen, S., Wu, J., Yin, S.: ARSpy: breaking location-based multi-player augmented reality application for user location tracking. IEEE Trans. Mob. Comput. **21**(2), 433–447 (2020)

33. Datta, A., Sen, S., Zick, Y.: Algorithmic transparency via quantitative input influence: theory and experiments with learning systems. In 2016 IEEE Symposium on Security and Privacy (SP), pp. 598–617. IEEE (2016)
34. Singh, M., Singh, S. K., Kumar, S., Madan, U., Maan, T.: Sustainable framework for metaverse security and privacy: opportunities and challenges. In: Nedjah, N., Martínez Pérez, G., Gupta, B.B. (eds.) International Conference on Cyber Security, Privacy and Networking (ICSPN 2022). ICSPN 2021. Lecture Notes in Networks and Systems, pp. 329–340 (2022). Cham, Springer International Publishing. https://doi.org/10.1007/978-3-031-22018-0_30
35. Nevelsteen, K.J.: Virtual world, defined from a technological perspective and applied to video games, mixed reality, and the metaverse. Comput. Anim. Virt. Worlds **29**(1), e1752 (2018)
36. Nguyen, C.T., Hoang, D.T., Nguyen, D.N., Dutkiewicz, E.: Metachain: a novel blockchain-based framework for metaverse applications. In: 2022 IEEE 95th Vehicular Technology Conference:(VTC2022-Spring), pp. 1–5. IEEE (2022)
37. Pooyandeh, M., Han, K.J., Sohn, I.: Cybersecurity in the AI-based metaverse: a survey. Appl. Sci. **12**(24), 12993 (2022)

Enhancing Educational Assessment: Leveraging Item Response Theory's Rasch Model

Georgi Krastev, Valentina Voinohovska(✉), and Vanya Dineva

University of Ruse, Ruse, Bulgaria
{georgk,vvoinohovska,vdineva}@uni-ruse.bg

Abstract. In the realm of educational assessment, accurate measurement of students' knowledge and abilities is crucial for effective teaching and learning. Traditional assessment methods often fall short in providing precise and meaningful insights into students' aptitudes. However, Item Response Theory (IRT), a psychometric framework, offers a powerful toolset to address these limitations. This article proposes an exploration of IRT's models and their potential to enhance educational assessment practices.

Keywords: Item Response Theory · Rasch Model · Educational Assessment

1 Item Response Theory (IRT)

Item Response Theory (IRT) is a psychometric framework that offers a sophisticated and powerful approach to educational assessment [1]. It provides a robust methodology for analyzing the relationship between individuals' responses to test items and their underlying abilities or traits [2].

At its core, IRT is based on a set of fundamental principles. One key principle is that the probability of a correct response to an item is dependent on the individual's ability and the item's characteristics. IRT models allow us to estimate and interpret these item characteristics, such as difficulty and discrimination, which provide valuable insights into the assessment process.

1.1 Key Principles and Foundations of Item Response Theory

Item Response Theory (IRT) is built upon a set of key principles and foundations that form the basis of this powerful psychometric framework [9]. Understanding these principles is essential for comprehending the underlying concepts and applications of IRT in educational assessment. The key principles and foundations of IRT are:

Relationship between Responses and Abilities

IRT revolves around the idea that the probability of a particular response to an item depends on the individual's underlying ability or trait being measured. Unlike classical test theory, which focuses on overall test scores, IRT delves into the item-level characteristics and their interaction with test-taker abilities.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 194–202, 2024.
https://doi.org/10.1007/978-3-031-53552-9_17

Item Characteristics

IRT acknowledges that each test item possesses its own set of characteristics that affect how it functions within an assessment. Two essential item characteristics are difficulty and discrimination. Difficulty refers to the level of skill or ability required to have a 50% chance of responding correctly to an item. Discrimination indicates how well an item differentiates between individuals with different levels of ability.

Item Response Functions

IRT models item responses using item response functions. These functions describe the probability of a correct response as a function of the item's characteristics (e.g., difficulty) and the individual's ability. The item response function allows for the estimation of the probability of success for any given ability level.

Person Parameters

IRT introduces the concept of person parameters, which represent individuals' latent abilities or traits being measured by the test. These person parameters are estimated based on the pattern of item responses and provide a more precise measurement of individual abilities compared to classical test scores.

Latent Trait Uni-dimensionality

IRT assumes that the latent trait being measured is unidimensional, meaning that the assessment is focused on measuring a single underlying ability or trait. This assumption allows for simpler modeling and interpretation of the test results.

Item Parameter Estimation

IRT models involve estimating item parameters, such as item difficulty and discrimination, from the observed item response data. Various estimation methods are employed, including maximum likelihood estimation (MLE) and Bayesian estimation, to obtain accurate estimates of item parameters.

Test Information

IRT provides a measure of test information, which indicates the precision of the measurement across the ability continuum. Test information allows for identifying the range of ability levels where the test is most informative and where the measurement is less precise.

By adhering to these key principles and foundations, IRT enables a deeper understanding of the relationship between item characteristics, individual abilities, and item responses. This understanding, in turn, facilitates more accurate measurement, enhances item quality, and supports evidence-based decision-making in educational assessment.

1.2 Types of IRT Models

Item Response Theory (IRT) encompasses different types of models that allow for a comprehensive analysis of the relationship between item responses and individuals' abilities. The main types of IRT models include [4]:

One-Parameter Models

One-parameter models assume that item difficulty is the only parameter that affects the probability of a correct response. The Rasch model and the one-parameter logistic model are commonly used examples of one-parameter models. These models provide a simple representation of the relationship between item responses and abilities.

Two-Parameter Models

Two-parameter models introduce an additional parameter, item discrimination, alongside item difficulty. Item discrimination indicates how well an item can differentiate between individuals with different levels of ability. The two-parameter logistic model and the Birnbaum's 2PL model are popular examples of two-parameter models. These models allow for a more nuanced understanding of item characteristics and their impact on the probability of a correct response.

Three-Parameter Models

Three-parameter models include item difficulty, item discrimination, and a guessing parameter. The guessing parameter accounts for the likelihood of guessing a correct response to an item, even with limited knowledge or ability. The three-parameter logistic model and the Lord's 3PL model are prominent examples of three-parameter models. These models are particularly useful when dealing with multiple-choice items where guessing can occur.

The choice of IRT model depends on the specific measurement context, the type of item responses, and the research or assessment objectives [3]. Each model offers unique insights into the relationship between item characteristics and individuals' abilities, enabling a more accurate and detailed analysis of educational assessment data [7, 8].

1.3 Advantages of IRT Over Classical Test Theory in Educational Assessment

Item Response Theory (IRT) offers several advantages over classical test theory (CTT) in educational assessment [6]. These advantages contribute to more precise and informative measurement of individuals' abilities and improved assessment practices:

Individualized Measurement: IRT provides individualized measurement by estimating individuals' latent abilities or traits. Unlike CTT, which focuses on overall test scores, IRT models estimate person parameters that represent individuals' abilities on a continuous scale. This allows for more accurate and tailored assessment of individual strengths and weaknesses.

Item-Level Analysis: IRT focuses on the item-level characteristics, such as item difficulty and discrimination. By estimating these parameters, IRT enables a detailed analysis of how each item functions and contributes to the overall assessment. This item-level analysis helps identify problematic items, improve item quality, and enhance the assessment's validity.

Differential Item Functioning (DIF) Detection: IRT facilitates the detection of differential item functioning (DIF), which occurs when an item behaves differently for different groups of test-takers. DIF analysis helps identify potential biases in the assessment process and ensures fairness by evaluating item performance across various groups (e.g., gender, ethnicity, language proficiency).

Improved Item Calibration: IRT allows for precise item calibration, estimating item difficulty and discrimination parameters based on the observed item response data. This calibration results in more accurate and comparable item measures across different assessments or administrations. It enables the creation of item banks and supports computerized adaptive testing, where items are selected based on the test-takers' estimated abilities.

Flexibility and Model Fit Assessment: IRT offers flexibility in model selection and fit assessment. Researchers can choose from various IRT models based on the measurement context and data characteristics. Additionally, IRT provides fit indices to evaluate how well the chosen model fits the data. This helps ensure that the model adequately represents the relationship between item responses and abilities.

Efficient Use of Test Items: IRT allows for efficient item usage in assessments. By estimating item parameters, IRT identifies the difficulty levels of items. This information helps in selecting appropriate items for specific ability levels, optimizing the assessment's measurement precision and reducing measurement error.

Multidimensional Measurement: IRT accommodates multidimensional measurement by modeling multiple latent traits simultaneously. This is particularly useful when assessing complex constructs that involve different dimensions or skills. Multidimensional IRT models provide a more comprehensive understanding of individuals' abilities across multiple domains.

2 Rasch Model

The Rasch model, a widely used model in Item Response Theory (IRT), provides a powerful framework for analyzing item responses and estimating individuals' abilities or traits [5]. Named after the Danish mathematician Georg Rasch, this model offers valuable insights into the relationship between item characteristics and individual abilities.

The most general form of test problem models according to the IRT can be described by:

$$P(z) = c + (1 - c)\frac{e^{-7.7a(z-b)}}{1 + e^{-1.7a(z-b)}} \tag{1}$$

where P(z) is the probability for an individual's correct answer, standing on a 'z' position of the score scale; a, b and c are the model variables, respectively named discrimination, difficulty and guessing parameter. The statement (1) is the so-called logistic distribution.

The difficulty and discrimination parameters have similar meaning as in the classical model but they are with different measurements. We will briefly clarify the contents of these three parameters of the model described by the function (1).

Discrimination (a). This parameter shows the extent of usefulness of the problem for discriminating the different levels on the score scale. It is proportional to the characteristic curve slope in the inflection point. The steeper the slope, the higher the discrimination is. For the one-parameter model the discrimination is a constant in all test problems.

Difficulty (b). This parameter sets the position of the characteristic curve with regard to the score scale. The more difficult the problem, the more shifted to the right the curve is. The parameter b value corresponds to the score scale point with correct answer probability of 0.5.

Guessing (c). This parameter influences the characteristic function asymptote and represents the probability for an individual on the lowest score position to give a correct answer of the problem. In the one-and two-parameter models the guessing parameter is zero.

In the one-parameter model the only estimated parameter is the difficulty. The discrimination is same for all problems, i.e. it is 1.0. The characteristic curve form is same for all problems, differing only with regard to its position on the horizontal axis. Test problems with equal percentage of correct responses have same difficulty estimates.

Figure 1 shows the resulting characteristic curve for the one-parameter model. The scores are normally transformed into z-numbers. The normally transformed scores are obtained by subtracting the average score from the individual score and dividing the result by the standard deviation. Upon the assumption for a normal score distribution the normal transformation results are within the $-3, 3$ interval. After that the characteristic curve is built. The graph is constructed with the help of the software environment "MATLAB".

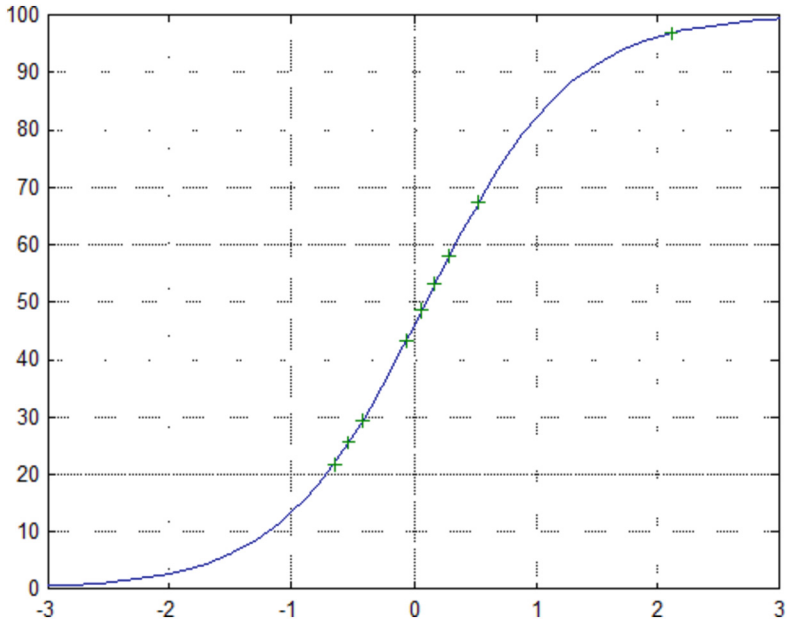


Fig. 1. Characteristic curve of one-parameter model.

Though the difficulty parameter is expressed just by a number, it is not a simple term. The goal of a test problem characteristic curve is to show the difficulties at different levels of the score scale. In result, a problem that is easy for individuals of high abilities can be difficult for individuals of lower abilities. That is why the difficulty parameter is implied only for individuals of equivalent abilities.

In order to establish the correspondence of the data to the theoretical logistic curve with relevant parameters, the Che-criterion is applied. It compares the differences between the observed data distribution and the assumed distribution along the logistic curve. The lower values correspond to more accurate correspondence between the curve and the data. It gets values within 0 and 1 and displays the probability of implementing the difference between the observed and assumed distribution in case that the problem is given to an analogous group of individuals. The results also show how well

the curve model form corresponds to the data of a problem. A lower Che-square statistics mean a good matching of the problem with the model.

The process of finding badly functioning problems in the models of IRT is completely different from that of the classical theory. Generally, it is based upon the researches related to the matching of the problem with the chosen model. The aids are not so simple.

3 Research and Results

The present paper includes a test on the course "Computer Networks and Communications" with a representative extract of 4th year students from the master's degree "Informatics and information technology in education" (Fig. 2).

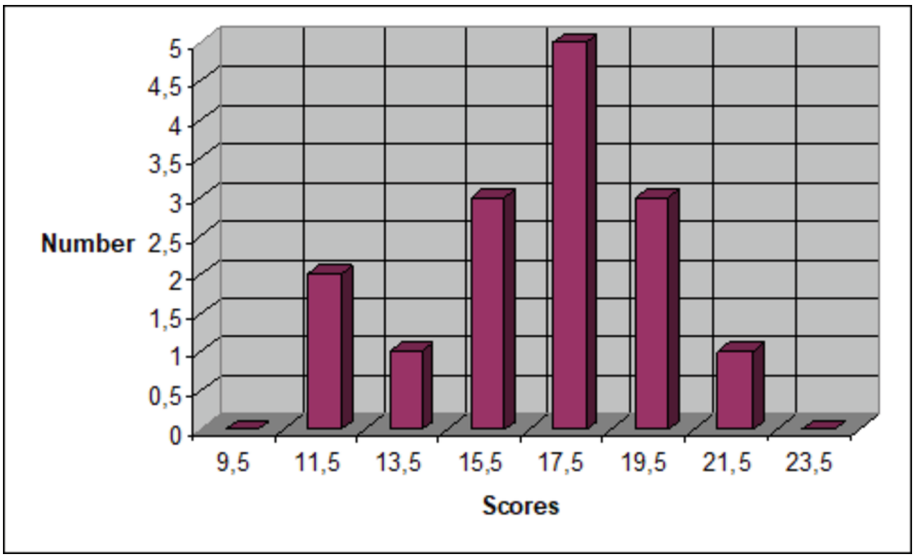


Fig. 2. Distribution of test scores for a representative extract.

The relative frequency of the measurements in the i-interval is obtained through dividing the interval frequency by the total number of the data (Fig. 3).

Calculated statistical parameters (characteristics) of the test:

- Average value 17,53.
- Mode 18.
- Median 17,5.
- Dispersion 6,16.
- Standard deviation 2,48.

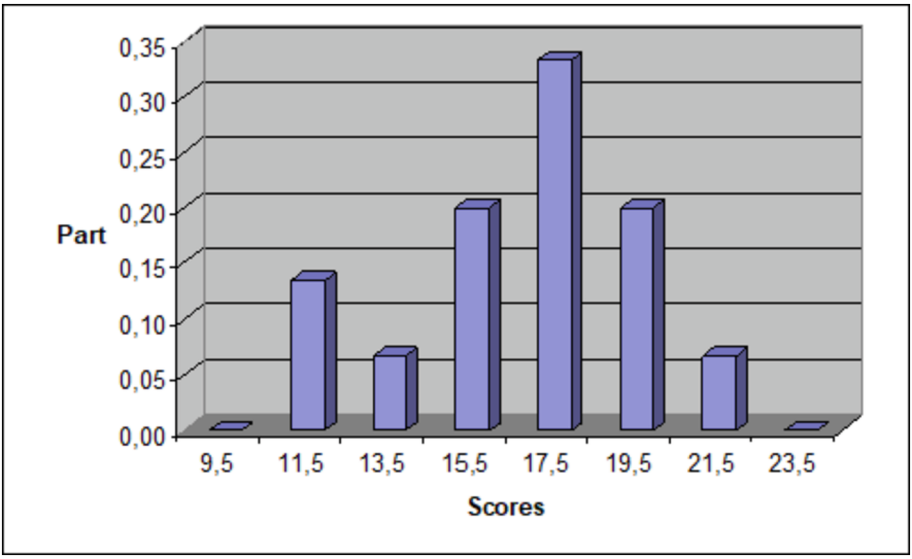


Fig. 3. Che-graph of relative frequencies.

The difficulty parameter of a test problem with a multiple-choice answer is equal to the percentage of students who solved the problem correctly (Fig. 4).

Estimated Test Problems Difficulty (Table 1)

Table 1. Estimated test problems difficulty.

Problems number	
Very easy	4
Easy	3
Optimal	7
Difficult	3
Very difficult	8
Test average difficulty	57,33

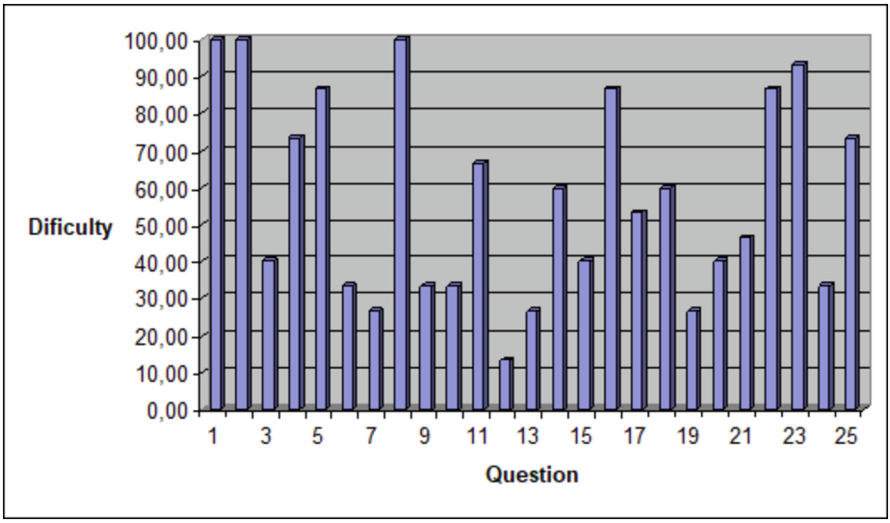


Fig. 4. Problems difficulty.

4 Conclusion

In conclusion, leveraging Item Response Theory's (IRT) one-parameter models presents a promising approach for enhancing educational assessment. This paper has explored the benefits and applications of IRT's Rasch Model in educational assessment, highlighting their ability to provide valuable insights into students' abilities and item characteristics.

By using IRT's Rasch Model, educators and assessment specialists can gain a deeper understanding of how students perform on specific test items. These models allow for the estimation of item difficulty and discrimination, enabling the identification of challenging items that may need revision or modification. Additionally, IRT's models provide a comprehensive framework for evaluating student abilities and estimating their proficiency levels, allowing for more accurate and meaningful assessment results.

Furthermore, the flexibility of IRT's models makes them applicable across various educational contexts and assessment types. Whether used in large-scale standardized testing or classroom-based formative assessments, these models can provide valuable insights into student performance, helping educators tailor instruction and interventions to meet individual learning needs.

The integration of IRT's models into educational assessment practices also holds the potential to enhance fairness and reduce bias. By focusing on item difficulty rather than relying solely on the overall test score, these models account for differences in item characteristics and ensure that students are not unfairly penalized or advantaged by certain items. This promotes a more equitable assessment environment and supports valid and reliable interpretations of students' abilities.

As educational assessment continues to evolve, leveraging IRT's models offers a powerful tool for improving the accuracy, fairness, and instructional relevance of assessments. However, it is important to acknowledge that the successful implementation of

these models requires expertise in test design, data analysis, and interpretation. Educators and assessment specialists should be equipped with the necessary knowledge and training to effectively utilize IRT's models and maximize their benefits in educational settings.

In conclusion, the application of IRT's models has the potential to transform educational assessment by providing a robust framework for understanding student abilities and item characteristics. By embracing these models, educators can enhance the quality of assessments, tailor instruction to individual needs, and promote equitable practices in education. Continued research, collaboration, and professional development in this field are crucial to fully harness the potential of IRT's models and advance the field of educational assessment.

Acknowledgments. This publication is developed with the support of Project BG05M2OP001-1.001-0004 UNITe, funded by the Operational Programme "Science and Education for Smart Growth", co-funded by the European Union trough the European Structural and Investment Funds.

References

1. Baker, B., Kim, H.: Item Response Theory: Parameter Estimation Techniques. CRC Press (2004)
2. De Ayala, J.: The Theory and Practice of Item Response Theory. The Guilford Press (2009)
3. Doncheva, J., Melandri, S., Valentini, M.: Technology as a school aid in motor activity. In: EDULEARN21 Proceedings, pp. 955–963 (2021)
4. Embretson, E.: Cognitive Design Systems in Item Response Theory: A Framework for Test Design. Routledge (2013)
5. Engelhard, G.: Invariant Measurement: Using Rasch Models in the Social, Behavioral, and Health Sciences. Routledge (2013)
6. Hambleton, K., Han, N.: Adapting Educational and Psychological Tests for Cross-Cultural Assessment. Routledge (2013)
7. Shoilekova, K.: Advantages of data mining for digital transformation of the educational system. In: Silhavy, R. (ed.) Artificial Intelligence in Intelligent Systems. LNNS, vol. 229, pp. 450–454. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-77445-5_42
8. Shoilekova, K.: Intelligent data analysis using a classification method for data mining knowledge discovery process. In: Silhavy, R. (ed.) Artificial Intelligence in Intelligent Systems. LNNS, vol. 229, pp. 153–157. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-77445-5_14
9. Von Davier, M.: Statistical Models for Test Equating, Scaling, and Linking. Springer, New York (2008). https://doi.org/10.1007/978-0-387-98138-3

Particular Analysis of Regression Effect Sizes Applied on Big Data Set

Tomas Barot[1(✉)], Marek Vaclavik[2], and Alena Seberova[3]

[1] Department of Mathematics with Didactics, Faculty of Education, University of Ostrava, Fr. Sramka 3, 709 00 Ostrava, Czech Republic
Tomas.Barot@osu.cz

[2] Department of Education and Adult Education, Faculty of Education, University of Ostrava, Fr. Sramka 3, 709 00 Ostrava, Czech Republic
Marek.Vaclavik@osu.cz

[3] Department of Preprimary and Primary Education, Faculty of Education, University of Ostrava, Fr. Sramka 3, 709 00 Ostrava, Czech Republic
Alena.Seberova@osu.cz

Abstract. In accordance with quantitative research, a wide spectrum of techniques can be seen. In the case of cardinal variables, regression analyses are suitable tools for the expression of dependences between observed variables. One of their options, the regression coefficients have been considered. However, the effect sizes analyses have not been so widely seen in research works in general. In this contribution, the big data analysis is being presented focusing on the regression effect size behavior following changing the number of samples. Two-dimensional and three-dimensional computations are applied with utilized mathematical regression models. The stable or stochastic behavior of Cohens f squared is discussed in the particular applied quantitative research of the OECD PISA with 397708 answers from respondents.

Keywords: Testing Dependences · Linear Regression · Effect Size · Educational Research · Big Data Set

1 Introduction

The guarantees about achieved results can not give so the stability of them in the quantitative research in accordance with the determined sample size.

The sample size can influent a wide spectrum of partial coefficients of obtained results, as can be seen in [1–3].

Besides the selection of an appropriate sample size, the testing normality [4] is necessary in the primary phase realized on data. However, the multiple or classical regression analyses [5–7] are not so assumed on this requirement.

In the case of the regression [5–7], the dependent and independent variables are considered. The linear behavior is obtained using the regression equation complemented by R correlation coefficient [5–7]. Also in the multiple regression, in the multidimensional

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 203–209, 2024.
https://doi.org/10.1007/978-3-031-53552-9_18

case [1, 5–7] with 1 dependent variable and m independent variable has the alternative R as multiple or adjusted correlation coefficient. Both types of source variables are classified as cardinal. Also, there can be imagined the binary type of variable for the special case of logistic regression [5–7].

Each regression equation can provide the prediction view for the purposes of estimating the behavior in the measured situation in general.

Especially, the multidimensional case brings special methods, which can be applied to multi-sample situations such as ANOVA, Kruskal-Wallis, multiple regression, etc. (with more than 3 samples, including) [8].

Each of these methods obtains the p value [6] as confirmation, of whether the declared hypothesis is failed to reject or reject in favor of the alternative hypothesis. Across the topic of regression, the hypothesis can be declared as "There are no linear dependences between the dependent and independent variable(s) on the significance level 5% or 1% or 0.1%.". While the significance level used to be determined as: 0.05 (social science) [9–12], 0.01 (technical science) [13] or 0.001 (medical science) [14, 15].

The concrete rejection of the zero hypothesis brings the question of how strong is the zero hypothesis rejected, if it is rejected in favour of the alternative hypothesis. For these purposes, there exist effect sizes for each of the considered methods.

In this paper, the guarantee of the regression parameters and effect size is discussed bound on the changing sample size N on the big data set of the OECD PISA research [16] with 397708 respondents across the world focus.

2 Exploration of Behaviour of Regression Effect Size

The synchronization of achieved results of statistical coefficients can be influenced by the sample size of the data. Considering sample size N can change the final results. A wide spectrum of those statistical parameters can have dynamic behavior as can be seen in [1–3]. On these specifications has been each quantitative research characterized.

In this paper, the method is presented as showing the influence between the correlation and final effect size - the Cohen f^2 in the frame of the regression analysis. The guarantee of stable behavior is discussed [8].

The regression function is assumed with the determination of the correlation coefficient R, in the case of multi-regression with m variables as multiple or adjusted R [5]. In the regression models, the dependencies of numerical cardinal variables are computed. For the case of two-dimensional measurement, the equation is (1), and for multidimensional case (2a)–(2b), when the Cohen f^2 is defined as (3). The regression coefficients are bounded with their lower indices with the order of the independent variables. The coefficient with lower index c is the regression constant. The dependent variable is denoted as y. All variables are cardinal variables [5–7].

$$y = \alpha_1 x_1 + \alpha_c \tag{1}$$

$$y = \alpha_1 x_1 + \ldots + \alpha_m x_m + \alpha_c \tag{2a}$$

$$y = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_c \tag{2b}$$

In general, the effect size can be an appropriate tool for clarification of the strongness of dependences. When the zero hypotheses are rejected in favor of the alternative hypotheses and have the distribution, the Cohen f^2 (3) can verify small $f^2 > 0.02$, medium >0.15, large >0.35 rejection of zero hypotheses [8].

$$f^2 = R^2 \cdot \left(1 - R^2\right)^{-1}; f^2 \in \mathbb{R} \tag{3}$$

The effect size and the (multiple) correlation coefficient can be displayed in dependences on the considered changing sample size N as can be seen in the following section.

3 Results

Table 1. Achieved Testing Dependences (1) in Accordance with Changing N

N	α_1	α_c	R	f^{22}
390000	0.011	6.792	0.073	0.005
360000	0.011	6.803	0.075	0.006
330000	0.011	6.788	0.077	0.006
300000	0.013	6.756	0.085	0.007
270000	0.012	6.782	0.082	0.007
240000	0.013	6.763	0.086	0.007
210000	0.012	6.755	0.082	0.007
180000	0.011	6.815	0.074	0.005
150000	0.010	6.849	0.069	0.005
120000	0.010	6.980	0.067	0.004
90000	0.011	6.913	0.077	0.006
60000	0.009	6.959	0.063	0.004
30000	0.008	7.101	0.044	0.002
20000	0.007	7.176	0.041	0.002
10000	0.003	7.840	0.019	0.000
5000	0.007	8.338	0.059	0.003
1000	0.005	8.459	0.043	0.002
500	0.003	8.387	0.027	0.001
100	0.025	7.537	0.221	0.052
50	0.024	7.525	0.255	0.069

Changing the sample size N is further discussed by the behavior of two and multi-dimensional (three) regression characteristics on the big data of the applied quantitative research. Concretely, the table and graphical comparisons are complemented by a correlation coefficient and the Cohen effect size.

In Excel, greater samples (to 300000) were considered for the case of two-dimensional regression, than in the case of three-dimensional regression computed in PAST Statistics 4.0 [17] (to 180 000).

In both cases, the dependent variable is y which means "*Well-being of each respondent*". And the independent variables are x_1 "*Age*" and x_2 "*Number of summarized taught hours by each of respondent*." of the PISA OECD research 2018 [16].

In Table 1, resulted coefficients of the two-dimensional case of (1) are obtained with regard to changing the sample sizes.

In Table 2, resulted coefficients of the three-dimensional case of (2b) are obtained with regard to changing the sample sizes.

Table 2. Achieved Testing Dependences (2b) in Accordance with Changing N

N	α_1	α_2	α_c	R	f^2
180000	-4.00×10^{-2}	0.010648	7.4462	0.073713	0.005
150000	−0.049384	0.010092	7.6295	0.069705	0.005
120000	−0.13525	0.009564	9.1183	0.068476	0.005
90000	−0.15553	0.011019	9.3718	0.078668	0.006
60000	−0.13825	0.009308	9.1424	6.47×10^{-2}	0.004
30000	−0.11974	0.007629	8.9890	0.045941	0.002
20000	−0.23032	0.006748	10.814	0.048039	0.002
10000	−0.34263	0.002882	13.249	0.044223	0.002
5000	−0.32236	0.007011	13.426	0.074561	0.006
1000	0.02811	0.005182	8.0143	0.044150	0.002
500	−0.2938	0.003444	13.029	0.049925	0.002
100	−0.54501	0.024601	16.190	0.236500	0.059
50	−0.47082	0.024306	14.902	0.520000	0.371

In the graphical representation in Fig. 1 and Fig. 2 with consideration of the scale of the vertical axes with N, the coefficients α_1 or α_2 are more stable for increased N.

For Fig. 3 and Fig. 4, the shapes of the displayed R and the Cohen f^2 seem to be approximately similar in the trend view for higher sample size N. Therefore, the interpretation of linear dependences can be not the same interpreted for a small sample size. Medians of R and the Cohen f^2 have been significant differenced across changing N; for sequence R (Table 1 and 2) and also for sequence f^2 (Table 1 and 2), with all cases

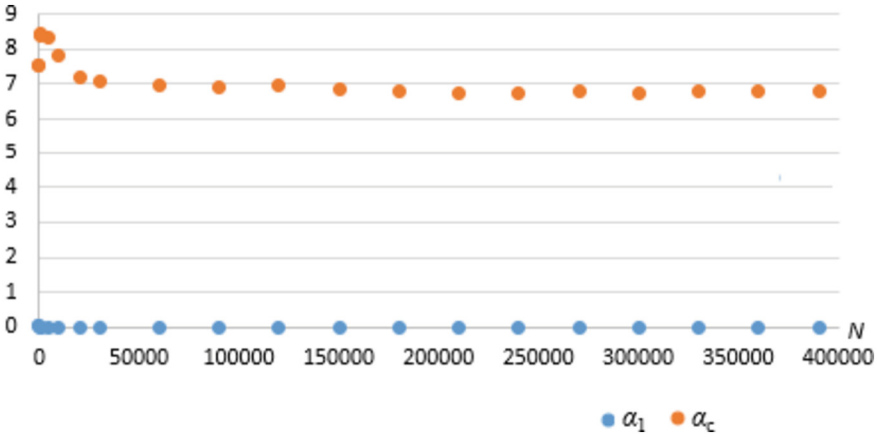


Fig. 1. Regression Coefficients α_1 and α_c of (1) in Accordance with Changing N

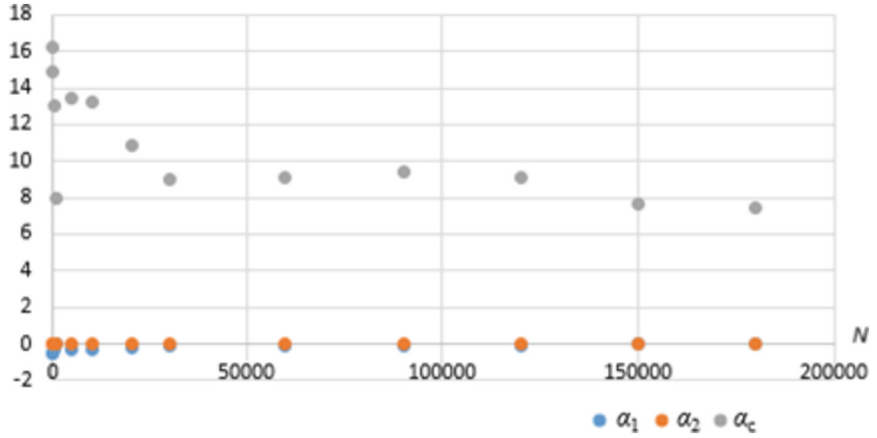


Fig. 2. Regression Coefficients α_1, α_2 and α_c of (2b) in Accordance with Changing N

of $p < 0.05$. These differences were measured by Wilcoxon one-sample test [5]. As can be seen, all statistical observed parameters are stable in the frame of higher sample size N.

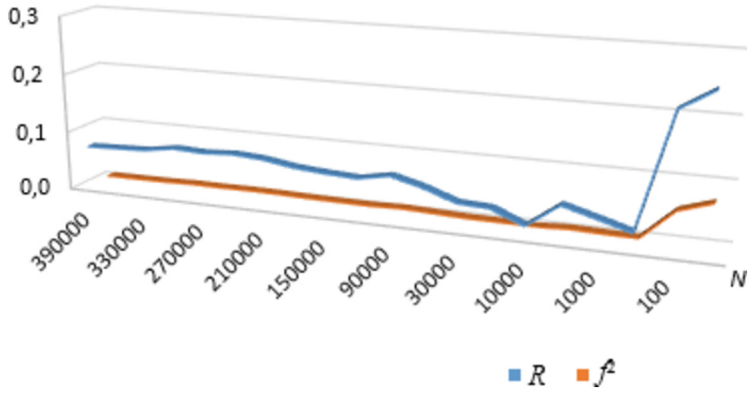


Fig. 3. Coefficients R and Effect Sizes f^2 of (1) in Accordance with Changing N

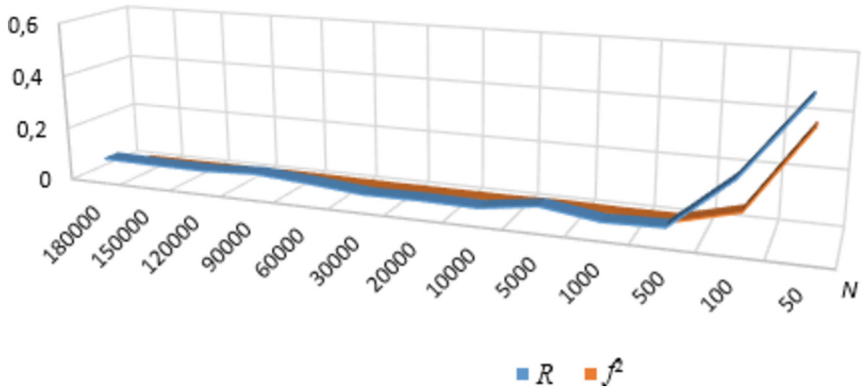


Fig. 4. Coefficients R and Effect Sizes f^2 of (2b) in Accordance with Changing N

4 Conclusion

In the discussed situation of the big data analysis with $N = 397708$, the regression parameters were observed regarding the changed sample size. In the case of two-dimensional and three-dimensional regression, the same effect can be seen, when the sample size is increased, the stability of the computations can be appropriately guaranteed. This analysis can be suitable for other researchers for satisfying the stability or guarantee of the definition of the sample size. Because, for different cases, the hypotheses can be in the first case confirmed and for other sample sizes rejected. Quantitative research with a big data set can be a suitable case for discussion of this situation. The particular research was selected from the OECD PISA research with a big amount of questionnaire answers and can be inspirative for other similar realized proposals of other researchers.

References

1. Vaclavik, M., Barot, T., Valisova, A., et al.: Analysis of cardinal-variables' dependences regarding models' structures in applied research of PISA. In: 2022 7th International Conference on Mathematics and Computers in Sciences and Industry: 7th International Conference on Mathematics and Computers in Sciences and Industry (MCSI), pp. 151–154. IEEE Computer Society Conference Publishing Services (2022)
2. Vaclavik, M., Sikorova, Z., Barot, T.: Skewness in applied analysis of normality. In: Silhavy, R., Silhavy, P., Prokopova, Z. (eds.) Software Engineering Perspectives in Intelligent Systems. AISC, vol. 1295, pp. 927–937. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-63319-6_86
3. Vaclavik, M., Sikorova, Z., Barot, T.: Particular analysis of normality of data in applied quantitative research. In: Silhavy, R., Silhavy, P., Prokopova, Z. (eds.) Computational and Statistical Methods in Intelligent Systems. AISC, vol. 859, pp. 353–365. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-00211-4_31
4. Tomsik, R.: Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov and Jarque-Bera tests. Scholars J. Res. Math. Comput. Sci. **3**(3), 238−243 (2019)
5. Kitchenham, B., Madeyski, L., Budgen, D., et al.: Robust statistical methods for empirical software engineering. Empirical Softw. Eng., 1–52 (2016)
6. Corrado, M.: Qualitative research methods and evidential reasoning. Philosophy Soc. Sci. **49**(5), 385−412 (2019). https://doi.org/10.1177/0048393119862858
7. Gauthier, T.D., Hawley, M.E.: Statistical methods. In: Introduction to Environmental Forensics, 3rd edn., pp. 99–148. Elsevier (2015). https://doi.org/10.1016/B978-0-12-404696-2.00005-9
8. Gupta, B.C., Guttman, I., Jayalath, K.P.: Statistics and Probability with Applications for Engineers and Scientists Using MINITAB, R and JMP. Wiley (2020)
9. Lackova, L.: Protective factors of university students. New Educ. Rev. **38**(4), 273–286 (2014). ISSN 1732-6729
10. Korenova, L.: GeoGebra in teaching of primary school mathematics. Int. J. Technol. Math. Educ. **24**(3), 155–160 (2017)
11. Schoftner, T., Traxler, P., Prieschl, W., Atzwanger, M.: E-learning introduction for students of the first semester in the form of an online seminar. In: Pre-Conference Workshop of the 14th E-Learning Conference for Computer Science, pp. 125−129. CEUR-WS (2016). ISSN 1613-0073
12. Sikorova, Z., Barot, T., Vaclavik, M., Cervenkova, I.: Czech university students' use of study resources in relation to the approaches to learning. New Educ. Rev. **56**(2), 114–123 (2019)
13. Barot, T., Krpec, R., Kubalcik, M.: Applied quadratic programming with principles of statistical paired tests. In: Silhavy, R., Silhavy, P., Prokopova, Z. (eds.) Computational Statistics and Mathematical Modeling Methods in Intelligent Systems. AISC, vol. 1047, pp. 278–287. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-31362-3_27. ISBN 978-3-030-31361-6
14. Sulovska, K., Belaskova, S., Adamek, M.: Gait patterns for crime fighting: statistical evaluation. Proc. SPIE – Int. Soc. Opt. Eng. **8901** (2013). ISBN 978-081949770-3. https://doi.org/10.1117/12.2033323
15. Svoboda, Z., Honzikova, L., Janura, M., et al.: Kinematic gait analysis in children with valgus deformity of the hindfoot. Acta Bioeng. Biomech. **16**(3), 89−93 (2014). ISSN 1509-409X
16. Organisation for Economic Co-operation and Development (OECD), 6 June 2023. www.oecd.org/pisa/
17. Hammer, O., Harper, D.A.T., Ryan, P.D.: PAST: paleontological statistics software package for education and data analysis. Palaeontol. Electronica **4**(1) (2001). http://palaeo-electronica.org/2001_1/past/issue1_01.htm

Applied Analysis of Differences by Cross-Correlation Functions

Tomas Barot[1(✉)], Ladislav Rudolf[2], and Marek Kubalcik[3]

[1] Department of Mathematics with Didactics, Faculty of Education, University of Ostrava, Fr. Sramka 3, 709 00 Ostrava, Czech Republic
Tomas.Barot@osu.cz

[2] Department of Technical and Vocational Education, Faculty of Education, University of Ostrava, Fr. Sramka 3, 709 00 Ostrava, Czech Republic
Ladislav.Rudolf@osu.cz

[3] Department of Process Control, Faculty of Applied Informatics, Tomas Bata University, Nad Stranemi 4511, 760 05 Zlin, Czech Republic
Kubalcik@fai.utb.cz

Abstract. The testing of the statistically significant similarities between two different data series cannot be measured using by Wilcoxon or T paired test, because there is not exists pretest and posttest. For the alternative way of testing the similarities, when data has some boundness, the proposal of using cross-correlation tests is presented in this paper. When necessary, the condition is existing some same characteristics, which can be seen as the connection of two power system nodes in the power transmission system. The cross-correlation analysis can be suitably complemented using by computing the p values between the correlation data themselves, which in addition, the trend of correlation data is measured for expression of the dynamic of cross-correlation achieved data. Application of the proposed method is realized on the real data set of the behavior of power transmission systems in the Moravia-Silesian region.

Keywords: Testing Differences · Cross-Correlation Function · p-value · Applied Research · Transmission Power System

1 Introduction

In the case of comparison, two samples comparison [1] cannot be compared by the standard statistical tools in the frame of paired testing [2–4]. The dynamics [5] of two processes based on time-series data sets can be suitably compared using the cross-correlation functions by their displaying as can be seen in this paper.

The guarantees about achieved results can be appropriately complemented by computing *p*-values [2–4], which can give information about internal behavior inside the time series data sets.

The practical application can be then implemented in case of time series between two processes. The applied research is concretely bound to the analysis of power transmission systems in the Moravia-Silesian region [6–8].

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 210–216, 2024.
https://doi.org/10.1007/978-3-031-53552-9_19

Line parameters have been instantly observed by expert teams in the frame of energy dispatching. The transmission grid system of this company (CEPS company in the Czech Republic) can save the measured data into the databases TRIS. Achieved results of the predictive information can be a suitable tool for experts regarding the appropriate reactions for purposes of the minimization of technical losses [6].

The specific area of the transmission system has been selected as nodes within a grid of 400 kV lines of the Moravian-Silesian region, which is specific by other weather dynamics as other places in the Czech Republic. Concretely, the following substations have been considered: Horni Zivotice, Dlouhe Strane, Kletna, Krasikov, Nosovice, and Prosenice. For these selected substations, variations of the transmitted power can be seen [6].

Particularly, the calculation of predictions using by the predictions utilities with the following evaluation of losses [6] that occur in the transmission system located in a certain part of the Czech Republic [6].

In this paper, the similarities of the dynamics trends between the two data sources are discussed bound on the cross-correlation analysis time data set of the CEPS data [9] of the power transmission system.

2 Applied Observing Differences by Cross-Correlations

Presented consequences have been analyzed by a statistical analysis of the tightness of considered measured and dynamics in addition to guaranteeing the statistical significance [2–4]. Regarding the statistical significance, paired comparisons have not been so possibly analyzed in the field of the power systems related to the quantitative approach by the Wilcoxon or T paired tests [2–4].

According to the principle of the statistical significance guarantee included in the frame of similarities testing [1], the cross-correlation functions [10–12] were utilized. For purposes of influences of the transmission power, these types of functions have been analyzed in the statistical software PAST Statistics 2.17 [13]. In Fig. 1, the transmission system can be seen focusing on the Moravia-Silesian region [9].

The CEPS dispatching [9] has continuously analyzed the measured variables of the transmission system and has been being evaluated proposed changes using the TRIS (Telemetrically and Control Information System) (Fig. 2). The option of the predictive computation was developed in the frame of the first author's practical cooperation within networking of CEPS company and the university. With the integration of the controlled power plant Dlouhe Strane, six substations (Fig. 2) have been built in the considered scheme in the Moravia-Silesian region [1].

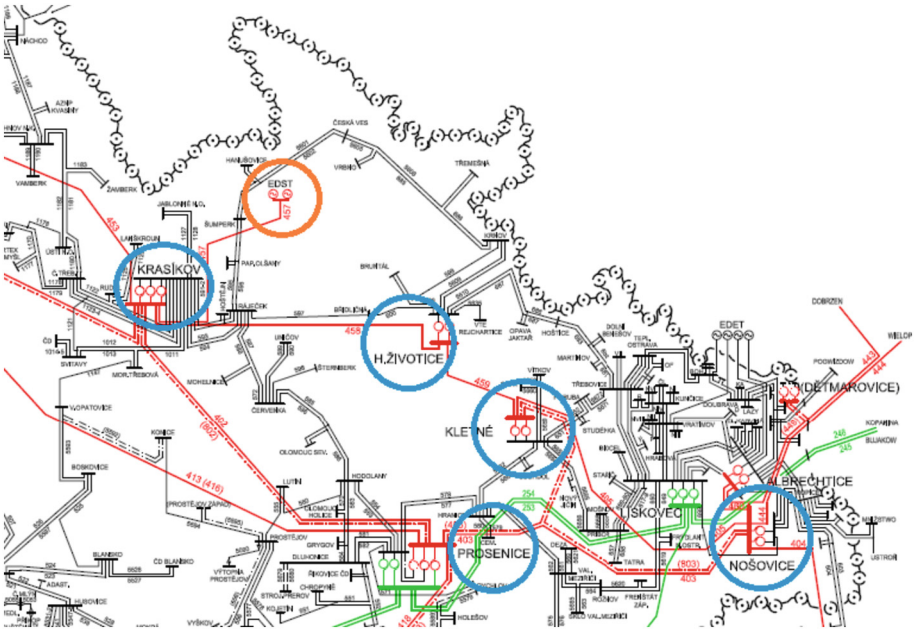


Fig. 1. Transmission Power Network Interactive Map CEPS in Detail Scale [9]

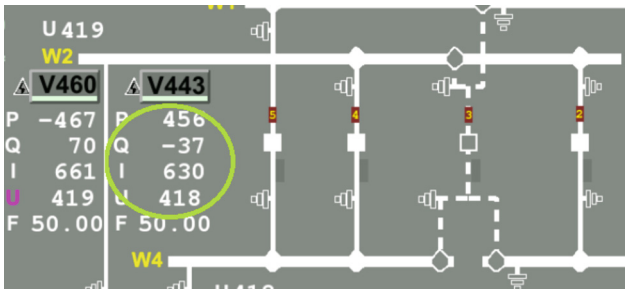


Fig. 2. CEPS's Interactive Panel of System TRIS at Dispatching Centre [9]

3 Results

Using the databases of TRIS system of company CEPS [9] the measured data and proposed regression mathematical models were considered.

The cross-correlation dependences [10–12] between behavior in given nodes (Dlouhe Strane DST, Krasikov KRA, Hladke Zivotice HZI, Kletne KLT, Prosenice PRN, Nosovice NOS) were tested in the software PAST Statistics 2.17 [13] and then displayed into the graphical representation. Concrete cross-correlation analysis can be seen: in Fig. 3 (PRN × NOS), in Fig. 4 (NOS × KLT), in Fig. 5 (KLT × HZI), in Fig. 6 (KRA × PRN), in Fig. 7 (KRA × DST). Where the p value has been confirmed the strongness of tightness.

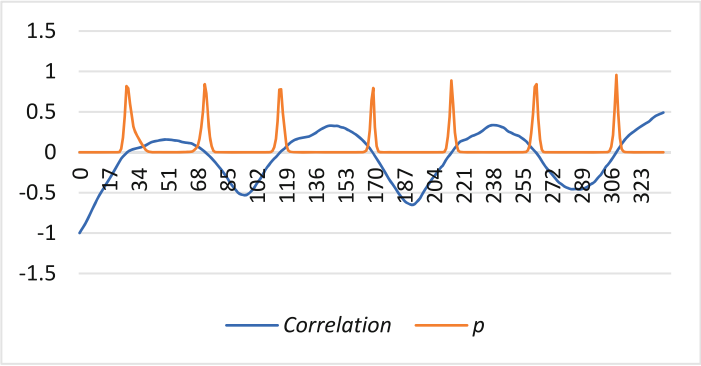


Fig. 3. Cross-Correlation Function between Nodes (PRN × NOS)

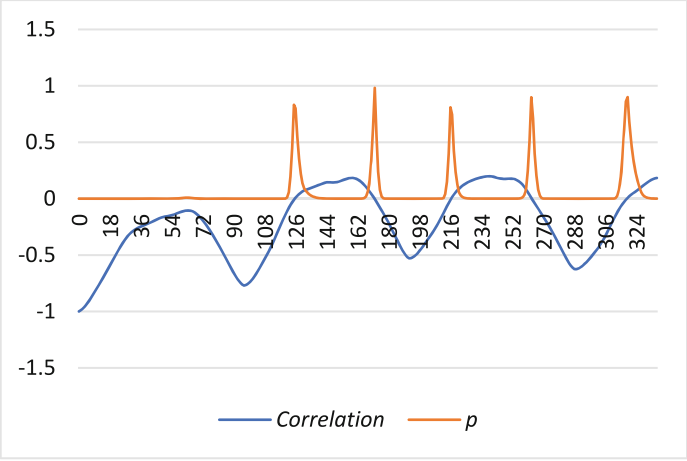


Fig. 4. Cross-Correlation Function between Nodes (NOS × KLT)

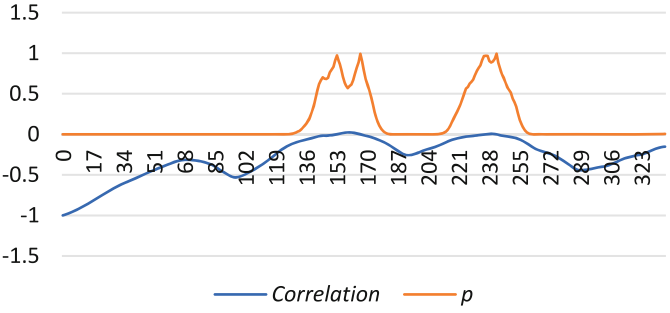


Fig. 5. Cross-Correlation Function between Nodes (KLT × HZI)

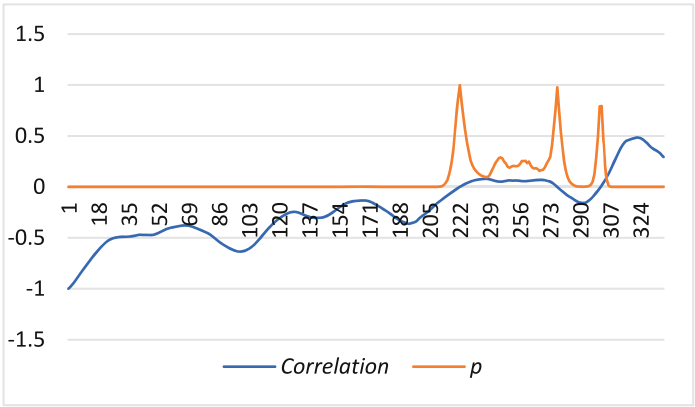


Fig. 6. Cross-Correlation Function between Nodes (KRA × PRN)

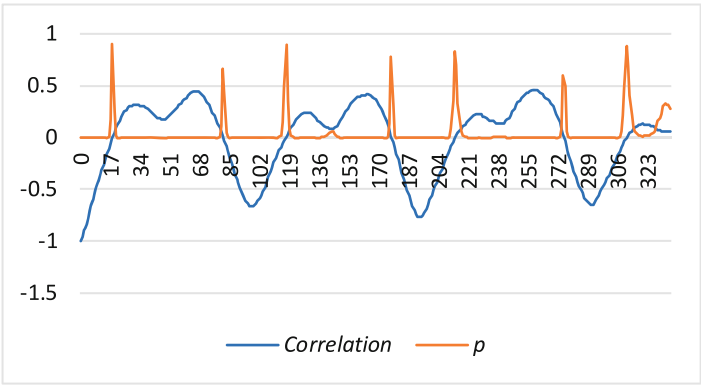


Fig. 7. Cross-Correlation Function between Nodes (KRA × DST)

An interesting comparison can be seen of combination the Fig. 3–4 (in Fig. 8) when the paired comparison Fig. 3 and 4 is being expressed the similarity with the tightness by both processes, where the NOS is the similar node of these 2 analyzed paths.

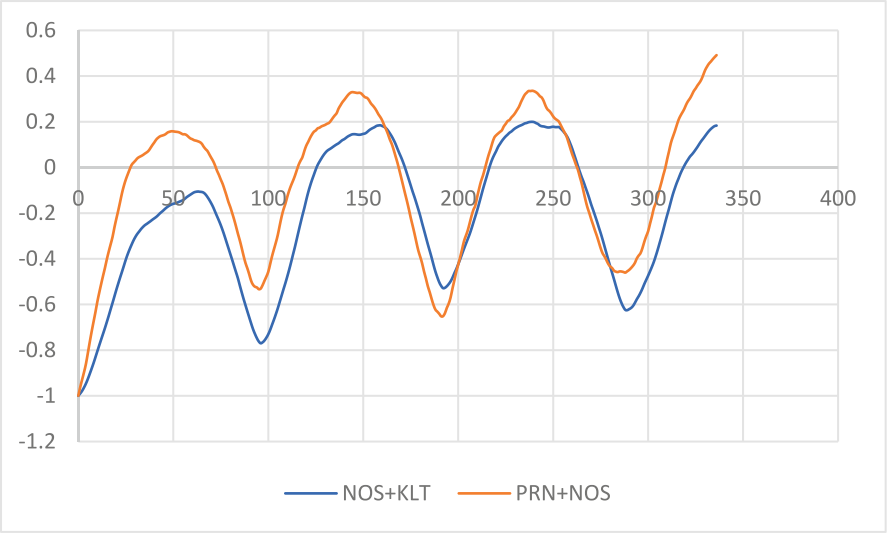


Fig. 8. Cross-Correlation Functions between Nodes PRN × NOS and NOS × KLT

4 Conclusion

In the paper, the consequences of the transmitted power have been supported by an additional statistical analysis of the cross-correlation tightness of considered measured variables in addition to guaranteeing the statistical significance. Regarding the statistical significance, paired comparisons have not been analyzed, although, the values of paired time series samples were compared in the field of the power systems related to the quantitative approach by the cross-correlations. Particularly, the signal processing is able to be giving feedback on the boundness effect of dynamics and similarities of obtained results. In the case of the transmission network node Nosovice, similar trends of the cross-correlation function were achieved. The comparison has confirmed the boundness between two paths with this considered node.

References

1. Barot, T., Krpec, R., Kubalcik, M.: Applied quadratic programming with principles of statistical paired tests. In: Silhavy, R., Silhavy, P., Prokopova, Z. (eds.) Computational Statistics and Mathematical Modeling Methods in Intelligent Systems. AISC, vol. 1047, pp. 278–287. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-31362-3_27
2. Kitchenham, B., Madeyski, L., Budgen, D., et al.: Robust statistical methods for empirical software engineering. Empirical Softw. Eng. **22**, 1–52 (2016)
3. Corrado, M.: Qualitative research methods and evidential reasoning. Philosophy Soc. Sci. **49**(5), 385−412 (2019). https://doi.org/10.1177/0048393119862858
4. Gauthier, T.D., Hawley, M.E.: Statistical methods. In: Introduction to Environmental Forensics, 3rd edn., pp. 99–148. Elsevier (2015). https://doi.org/10.1016/B978-0-12-404696-2.00005-9

5. Barot, T., Lackova, L.: Confirmation of cybernetical modelling in applied research of resilience. In: Silhavy, R. (ed.) Informatics and Cybernetics in Intelligent Systems. LNNS, vol. 228, pp. 255–263. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-77448-6_24
6. Rudolf, L., Barot, T., Bernat, M., et al.: Influence of temperature and transmitted power on losses in particular transmission system. WSEAS Trans. Power Syst. **17**, 53–61 (2022). ISSN 1790-5060
7. Dong, Y., Shi, Y.: Analysis of losses in cables and transformers with unbalanced load and current harmonics. J. Electric. Electron. Eng. **9**(3), 78–86 (2021)
8. Ruppert, M., Slednev, V., Finck, R., Ardone, A., Fichtner, W.: Utilising distributed flexibilities in the European transmission grid. In: Bertsch, V., Ardone, A., Suriyah, M., Fichtner, W., Leibfried, T., Heuveline, V. (eds.) Advances in Energy System Optimization. TM, pp. 81–101. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-32157-4_6
9. CEPS a. s.: Transmission Network in the Czech Republic and Central Europe in 2013–2015 in the context of EWIS. https://www.ceps.cz/en/studies-and-analyses. Accessed 9 June 2023
10. Fan, L., Chen, W., Jiang, X.: Cross-correlation fusion graph convolution-based object tracking. Symmetry **15**, 771 (2023). https://doi.org/10.3390/sym15030771
11. Mølvig, B.H., Bæk, T., Ji, J., Bøggild, P., Lange, S.J., Jepsen, P.U.: Terahertz cross-correlation spectroscopy and imaging of large-area graphene. Sensors **23**(6), 3297 (2023). https://doi.org/10.3390/s23063297
12. Ahmed Shahin, A., et al.: Maximum power point tracking using cross-correlation algorithm for PV system. Sustain. Energy Grids Netw. **34** (2023). ISSN 2352-4677
13. Hammer, O., Harper, D.A.T., Ryan, P.D.: PAST: paleontological statistics software package for education and data analysis. Palaeontol. Electronica **4**(1) (2001). http://palaeo-electronica.org/2001_1/past/issue1_01.htm

Paraphrasing in the System of Automatic Solution of Planimetric Problems

Sergey S. Kurbatov(✉)

Research Centre of Electronic Computing, Moscow 117587, Russia
curbatow.serg@yandex.ru

Abstract. An approach for synthesizing natural language sentences from ontological representation is developed. The approach is based on the method of semantically oriented paraphrasing. Preliminary results of computer realization of the approach in the subject area of planimetric problem solving are obtained. Experimental results in the aspect of synthesizing descriptions for inductive and plausible reasoning are discussed. Visualization tools (JSXGraph) provide a high visualization of synthesis and paraphrasing. In the applied aspect, the approach focuses on the creation of training systems that integrate the capabilities of linguistic processing of the problem text, automatic solution and interactive visualization. #COMESYSO1120.

Keywords: Semantic Paraphrasing · Geometric Problem Solver · Interactive Visualization · JSXGraph JavaScript Library

1 Introduction

The problem of paraphrasing in the aspect of computer implementation has been studied for decades by both linguists and developers of natural language computer communication systems. As early as 1970, the famous linguist Melchuk [1] created the concept of "Meaning ⇔ Text" focused both on the mapping of the natural language (NL) text into a semantic representation and on the synthesis of the text from this representation. The "Meaning ⇔ Text" model was to provide a compression of multiple synonymous NL descriptions into a single semantic structure, as well as the generation of a variety of texts with a fixed meaning.

The results of the research on the problem were used in machine translation (Apresyan [2]), in the development of the NL-interfaces for expert systems and database access (Popov [3], a book ahead of its time). The problem turned out to be quite complicated, and today the successes achieved are rather modest, especially if we consider the amount of effort spent by researchers and developers on its solution. At present, paraphrasing based on neural networks is widely used for applied purposes: translation, text generation (from advertising to scientific articles), anti-plagiarism, search in large arrays of texts, etc.

So far, successful NL interfaces can be implemented only in a given subject area and significant limitations on the language of the interface. The proposed research is

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 217–225, 2024.
https://doi.org/10.1007/978-3-031-53552-9_20

also focused on the semantics of the subject area (geometry) and the corresponding communication language. Nevertheless, the relevance of the research is determined by the following:

- the generality of the mechanism with paraphrasing rules in both analysis and synthesis of the NL;
- the semantic representation developed for the subject area, providing the possibilities of logical inference;
- the use of ontology to support paraphrasing and dialogue;
- logical and inductive reasoning for paraphrasing;
- interactive visualization of system actions;
- applied orientation to educational goals.

The generality of the mechanism is based on a standard transformation of syntactic structures, which is in principle independent of both the subject domain and partly of the language. The semantic representation contains references to standard syntactic structures in addition to logical inference, which makes it possible to synthesize different NL descriptions through paraphrasing.

The ontology ensures the generality of paraphrasing rules because the objects, properties and relations included in the rules are described on a high conceptual level if possible. Both syntactic structures and drawings as well as arbitrary Latex-style mathematical formulas are visualized using JSXGraph tools [4]. By clicking on the visualized objects it is possible to get detailed information about them. The study was performed on the basis of the system of automatic solution of planimetric problems [5, 6].

2 Methodology

There The methods used in this work are based on classical linguistic approaches (analysis and synthesis of the NL [3], meaning-text model [2]) and cognitive aspects of well-formalized problems [5]. Knowledge representation methods are used in an ontology that integrates knowledge about the NL, the subject domain, and visualization tools.

Paraphrasing is a rather ambiguous term and can be interpreted as:

- explaining or clarifying some text;
- clarification of the context of a statement;
- preserving the basic meaning of the paraphrased text.

The paraphrase in the proposed article is focused on NL interfaces to intelligent systems. Virtually all developers of modern NL interfaces use a combination of approaches (adapted sense-text model, syntactic templates, knowledge of tables, attributes and their values in database, reliance on semantics of subject area, menu elements in dialogs for complex cases, etc.). The proposed research is no different in this.

The syntactic structure formed by the linguistic translator contains word forms of the source text and references to morphological classes in the vertices [7]. The links between the word forms are maximally semantically loaded. The links are named with natural language question words (what?, when?, where?, what?, etc.), which significantly increases the visibility of the structure and defines a generally meaningful semantics (time, place, attribute, etc.).

The ontology contains a hierarchy of concepts associated with word forms, morphological classes and criteria for establishing syntactic relations between pairs of word forms (or between noun groups). Paraphrasing, as noted above, is performed over the formed syntactic structure.

Philosophically, paraphrasing is a very complex concept based on Frege's triangle "sign - meaning - significance". This study focuses on the algorithmic and computer development of paraphrasing. In the Frege's triangle aspect, our goal is to identify the meaning (program code), which allows to obtain the desired object-drawing (denotation) from the NL description (sign).

For example, the descriptions "the sum of the numbers seven and five" and "7 + 5" have the same denotation (number 2), but different meanings. Paraphrasing is goal-oriented; it should reveal a meaning that the computer can understand. In this case, the meaning is given by the semantic structures of the subject domain represented in the ontology. It is to this description that paraphrases should reduce some fragment of arbitrary text describing a task.

At the abstract level, the mechanism of paraphrasing as a synonymic transformation of tree structures is very general. For example, such a transformation can be treated as a synthesis of an SQL-program from the syntactic structure of the corresponding NL description. The abstract level implies that the left part of the paraphrasing rule contains concepts that are signified by concrete word forms of the syntactic structure. The generality of the mechanism is discussed in more detail below.

Cognitive structure in the broad sense is associated with the system of human cognition. In this robot we will understand cognitive structures as ontological structures used to formalize some area of knowledge by means of a conceptual scheme. As part of the computer implementation, the emphasis is on verbalizability, algorithmic and efficiency in solving a practical problem. Note that the model of the famous researcher on cognitive semantics Talmy [8] takes into account syntactic structures as cognitive as well.

3 Experiments Results

3.1 Paraphrasing and Syntax

In the course of the experiments, the texts of several hundred geometric problems of the possibility of their paraphrasing were analyzed. When testing the developed approach to paraphrasing, syntactic structures generated by the linguistic processor as well as structures prepared manually or generated from the semantic representation were used.

To visualize the results of paraphrasing, the synthesis of the NL by the current syntactic structure is implemented. The synthesis algorithm is recursive, starting from the "leaves" of the tree, the chain of the NL to the "ancestor" with alternatives is formed, then the "ancestors" with fully formed alternatives are selected. Then these "ancestors" are given the status of "leaves" and the process is repeated. It is beyond the scope of this article to discuss the details of the synthesis.

Let us explain the results with examples. Figure 1 shows the English and Russian text of a geometric problem, as well as the syntactic structure of the text (naturally, the Russian text).

In triangle ABC, point D is arbitrarily taken at the base of BC or at its continuation, and circles are circumscribed around triangles ACD and BAD.
Prove that the ratio of the radii of these circles is constant.
Find the position of point D for which these radii have the smallest value.

В треугольнике ABC на основании BC или на его продолжении взята произвольным образом точка D и около треугольников ACD и BAD описаны окружности.
Доказать, что отношение радиусов этих окружностей есть величина постоянная.
Найти такое положение точки D, для которого эти радиусы будут иметь наименьшую величину.

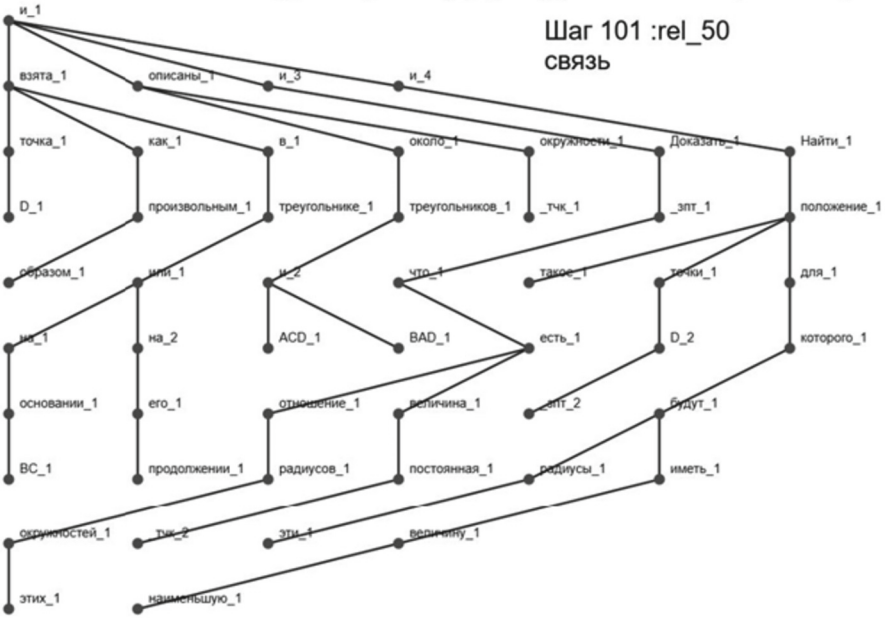


Fig. 1. The text of the problem in English, the Russian text and the actual syntactic structure.

The synthesis on the structure of Fig. 1 forms a text synonymous to the original text with the accuracy of permutations of name groups. For example, the beginning text can "have the form "Point D is taken arbitrarily in the triangle ABC at the base of BC or at its continuation…". As already emphasized, the goal of paraphrasing is to produce a text that is as oriented as possible semantic representation. This is what is further used by the system to construct the drawing and automatically solve the problem.

The semantics of the problem is described with the objects "point", "segment", "circle", "line"… and the relations "on", "start", etc. The description for the drawing alone contains about 10 triads of the "object1 relation otobject2" kind. A few more triads describe the actual conditions of the problem (relations, radii, minimum value). Let us explain the paraphrasing style on a simpler example.

Let the original text be "AD is the median of triangle ABC." Then the paraphrase, directly mapped to semantics, has the form: "Triangle ABC. Point D lies on the segment BC. The segment AD. BD = CD." The text of the paraphrase is synthesized from the syntactic structure, we won't stipulate it further on.

Returning to the structure in Fig. 1, we note that the presence of “its” and “these” references in the text requires their resolution. This is a special problem, but if it is solved, paraphrasing will provide a corresponding modification of the text: “its” → “the base BC”, “these circles” → “the circles are circumscribed around triangles ACD and BAD”, etc. Of course, this will make the text more cumbersome, but it simplifies the formation of the semantic representation.

Paraphrasing is also important for reconstructing incomplete (elliptical) sentences. Here are two examples of such paraphrases. “The diameter of the first circle is 12 cm, and the second circle is 4 cm”. The paraphrase of an incomplete sentence is formed by substituting objects into the complete syntactic structure, taking into account their ontological characteristics: “The diameter of the second circle is 4 cm”.

Similarly, by “The point A is inside the square and the segment BC is outside” the complete phrase “the segment BC is outside the square” is formed. The issue of ellipticity resolution is not considered in this paper, this issue is analyzed in more detail in [9]. Let us only note that interactive visualization provides meaningful paraphrase comments on ellipsis resolution.

The interactive visualization in the above example emphasizes the similarity of syntactic structures and the rationale of word-form substitution when resolving ellipticity. In the ontology, “point” and “segment” refer to the concept “#geometric_figure” and “inside” and “outside” to the concept “#localization”. This is what ensures the formation of the complete phrase. The symbol # means that it is a concept and not a real word form of natural language (“point,” “segment”).

We emphasize that the linguistic problem of resolving ellipticity and anaphoric references is quite complex. However, semantically-oriented structures and an ontology that takes into account both linguistic and subject knowledge form, in the author’s opinion, a solid foundation for solving this problem.

3.2 An Expanded View of Paraphrasing

In the ontology, paraphrasing rules are represented as structures in which objects and relations are concepts. The broader context of problem solving required us to take into account Poya’s method [10], related to plausible reasoning, when testing. For example, in a neatly made drawing, we note the parallelism of some lines. The Poya method proposes to use this as a working hypothesis, which it is reasonable to prove.

In this case, a plausible reasoning might be formulated as follows: “The drawing suggests that these lines are perpendicular. Try to prove that the angle between these lines is 90°.” The first sentence of the reasoning can have different formulations. For example:

You can see from the drawing that the lines are parallel.
Note that in the drawing the lines are almost parallel.
The straight lines in the drawing appear to be parallel.
The drawing suggests that the lines are parallel.
The drawing shows that the lines are parallel.
In the drawing the lines look parallel.
The drawing demonstrates the parallelism of the straight lines.
It looks like the lines in the drawing are parallel.

The drawing suggests that the straight lines are parallel.

The drawing allows us to speculate about the parallelism of the lines.

etc.

By paraphrasing, they are reduced to the canonical form of a rule in which the left-hand side has an ontological representation (simplified):

{#empirical #data} {#inferring} {#that} {#hypothesis}

The fact that the objects included in the rule are exactly concepts (though they have the form of word forms) is emphasized by the symbol #. The concepts are signified by the real word forms of the source text, and the right-hand side is given to the system solver:

{#prove} {#hypothesis}

In this case, the paraphrasing is done externally in a standard way, but the content is different. The right part is not synonymous with the left part and its formation models not logical but plausible reasoning. The given examples, of course, are not found in the texts of geometrical problems and were tested precisely for the extended view of paraphrasing.

Paraphrasing from semantic representation structures is based on the NL descriptions of triads. This allows from the description "Point A is the beginning of segment AB. Point B is the beginning of segment BC. Point C is the end of segment AC." by paraphrasing to the description "Triangle ABC". Similarly, from the detailed semantic description the descriptions of the form "AD - bisector/height/median of triangle ABC" are formed. The details of the semantic description (equality of angles/segments, perpendicularity) are omitted.

Let us briefly formulate the results:

- a paraphrasing approach that integrates elements of syntax semantics, domain ontology, and interactive visualization is developed;
- algorithms of the approach are developed;
- computer implementation of the algorithms was performed;
- testing of the implementation in a specific subject area has been carried out.

4 Discussion

The examples in Sect. 2 demonstrate the use of the paraphrasing apparatus for transforming syntactic structures and the subsequent synthesis of EL descriptions, providing clear results. The paraphrasing of both synonymous structures and structures with ellipses and references is demonstrated. It was also noted there that, in a broad sense, paraphrasing can be treated as a synthesis of an SQL program on the basis of a syntactic structure.

For example, based on the syntactic structure of the EA-query "List goods in delivery N5", an SQL-text was synthesized using paraphrasing for experimental purposes:

SELECT GOODS.name FROM GOODS, CAPTION, SPECIFICATION WHERE CAPTION.name = "N5" and CAPTION.cpID = SPECIFICATION.spID and SPECIFICATION.goodID = GOODS.goodID. The structure of this SQL-query obtained by paraphrasing the US text is shown in Fig. 2.

Of course, this is only a demonstration of the potential of the developed rephrasing. Nevertheless, it required the introduction of minimal information about relational tables (goods, caption, specification) and attributes into the ontology. In the developed version, with the help of paraphrases and interactive visualization, it is necessary to explain to the user about the relevant tables.

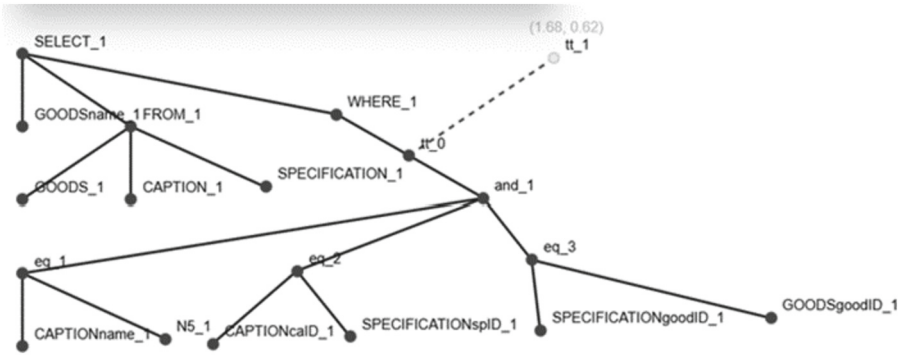


Fig. 2. Structure of the SQL-query (tt_0 - tt_1 - call of the interactive visualization by mouse click on the drawing object).

At the present time, the EJ interfaces to relational databases are not widespread, but the explanatory methods developed in this direction are quite interesting. In particular, their use for Explainable Artificial Intelligence (XAI) is very promising, because users of neural networks are concerned with the problem of justification and explanation of solutions recommended by the network.

The style of explanation should be determined by the ontology. For example, the ontology should provide explanations not at the level of "Resolving two Horn disjuncts generates NIL", but at the level of "statements X is different from Y and X matches Y are contradictory".

Because of the generality of the paraphrasing algorithm, the difference between synonymous transformations of syntactic structures and logical inference can be erased. In this case, when applying the paraphrasing rule, the comment-justification may look like: by Pythagoras' theorem, "by definition of a triangle", "by properties of triangle angles", "by definition". Example:

- The segment from the vertex of the triangle forms an angle with the base equal to the larger angle of *the triangle with sides 3, 4, and 5*;
- The segment from the vertex of the triangle forms an angle with the base equal to the greater angle of *the right triangle*;
- The segment from the vertex of the triangle forms an angle with the base *equal to 90°*;
- The segment from the vertex of the triangle *is the altitude*.

Paraphrased passages are italicized and underlined.

Some quantitative considerations regarding paraphrasing. In the development of the NL-interfaces, the author asked linguists to estimate the number of high-level paraphrasing rules for the whole NL. High level means that the rules do not include specific word forms, but concepts. Linguists' suggested estimate: 3,500 high-level rules will provide a reduction to some standard representation (meaning-text style) of the vast majority of NL texts.

A few dozens of subject-specific paraphrasing cases can be entered manually. The question of automatic detection of high-level rules is still a subject of research.

Modern technical means provide possibilities for storing knowledge, which in the early stages of computer processing of the NL seemed unattainable. In particular, the set of all word forms, triads of the type "morphological class 1 syntactic relation morphological class 2" and the set of all universal paraphrases together gives an inflated estimate of several millions.

The estimate for the Russian language is based on the number of bases of about 100,000 [11], the number of morphological classes and syntactic relations of 600 and a few dozen, respectively. The process of reducing an arbitrary text to a canonical universal representation (in the style of the sense-text model) is expedient to organize by means of parallel computer processing. All the more so because this process is performed on syntactic structures rather than on the text.

Related research [12, 13] are based on GSM8K, a set of NL descriptions and solutions of mathematical problems for schoolchildren. The research uses the methodology of forming a set of candidate solutions. Then a metric is used to select the best solution. The research is focused on algebraic problems without visualization.

Currently, commercial paraphrasing systems, usually based on neural networks (Robotext, QuillBot, WordAi, Spin Rewriter, SpinBot, etc.) are widely used. The main goals of such systems are:

- help in creating content and improving it;
- anti-plagiarism;
- fake news identification;
- business name generators;
- help in designing websites with AI etc.

Obviously, these goals are markedly different from the ontology-oriented paraphrasing investigated in this article. Nevertheless, even within this direction, attempts are being made to use neural network and ontology in hybrid applications (e.g., in medicine [14]), but so far they face serious problems. The hybrid approach uses a fuzzy logic reasoning mechanism [15], with ontology as a separate level of reasoning. Unfortunately, ontology has failed to improve the quality of fuzzy component inference.

5 Conclusion

Currently, paraphrasing based on neural networks is widespread. While recognizing the impressive successes of this approach, we note that the potential of classical linguistic paraphrasing for the purpose of extracting deep semantic content from a text is far from exhausted.

Within In particular, the use of ontology-based paraphrasing to describe the parameters of a neural network and its results opens the possibility to build rules explaining neural network decisions. Such capabilities are important for explicable artificial intelligence – XAI [16, 17], which allows users to understand how exactly a neural network arrived at its conclusions during data processing.

References

1. Mel'čuk, I.A.: Anna Wierzbicka, semantic decomposition, and the meaning-text approach. Russ. J. Linguist. **22**(3), 521–538 (2018)
2. Apresyan, J.D., Boguslavsky, I., Iomdin, L., Sannikov, V.: Theoretical problem of Russian syntax; interaction of grammar and vocabulary. Languages of Slavonic Cultures, Moscow, Russia (2010)
3. Popov, E.V.: Talking with Computers in Natural Language. Springer, Heidelberg (2011)
4. JSXGraph. https://jsxgraph.uni-bayreuth.de/share/. Accessed 6 June 2023
5. Kurbatov, S., Fominykh, I., Vorobyev, A.: Cognitive patterns for semantic presentation of natural-language descriptions of well-formalizable problems. In: Kovalev, S.M., Kuznetsov, S.O., Panov, A.I. (eds.) Artificial Intelligence. LNAI, vol. 12948, pp. 317–330. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-86855-0_22
6. Naidenova, X., Kurbatov, S., Ganapolsky, V.: Cognitive models in planimetric task text processing. Int. J. Cogn. Res. Sci. Eng. Educ. **8**(1), 25–35 (2020)
7. Kurbatov, S.: Linguistic processor of the integrated system for solving planimetric problems. Int. J. Knowl. Based Intell. Eng. Syst. **15**(4), 14 (2021). https://doi.org/10.4018/IJCINI.20211001.oa37
8. Talmy, L.: Semantics and Syntax of Motion. Academic Press (1975). https://doi.org/10.1163/9789004368828_008
9. Naidenova, X.: Cognitive processes in generating and restoring elliptical sentences. In: Proceedings of the Conference DAMDID-2021, Moscow (2022)
10. Polya, G.: Mathematik und plausibles Schliessen. Birkhäuser, Basel (1988)
11. Zaliznyak, A.: Grammatical Dictionary of the Russian Language. http://www.gramdict.ru/. Accessed 6 June 2023
12. Cobbe, K., et al.: Training Verifiers to Solve Math Word Problems (2021). https://doi.org/10.48550/arXiv.2110.14168
13. Golovneva, O., et al: ROSCOE: a suite of metrics for scoring step-by-step reasoning. In: Published as a Conference Paper at ICLR (2023). https://doi.org/10.48550/arXiv.2212.07919
14. van der Velden, B.H.M., Kuijf, H.J., Gilhuijs, K.G.A., Viergever, M.A.: Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. Med. Image Anal. **79**, 102470 (2022). https://doi.org/10.1016/j.media.2022.102470
15. Averkin, A., Yaruhev, S.: Review of research in the development of methods for extracting rules from artificial neural networks Izvestiya RAS. Theory Control Syst. **6**, 106–121 (2021)
16. Pradeepta, M.: Practical Explainable AI Using Python: Artificial Intelligence Model Explanations Using Python-based Libraries. Apress Media LLC (2022)
17. Hu, B., et al.: XAITK: the explainable AI toolkit. Appl. AI Lett. **2**(4), e40 (2021). https://doi.org/10.1002/ail2.40

The Formation of Ethno-Cultural Competence in Students Through the Use of Electronic Educational Resources

Stepanida Dmitrieva[1], Tuara Evdokarova[1], Saiyna Ivanova[1(✉)], Ekaterina Shestakova[2], Marianna Tolkacheva[3], and Maria Andreeva[3]

[1] Department of Social Pedagogy, Pedagogical Institute, M. K. Ammosov North-Eastern Federal University, NEFU, Yakutsk, Russia
saina.ivanova.99@mail.ru
[2] S. F. Gogolev Yakut Pedagogical College, Yakutsk, Russia
[3] V. M. Chlenov Yakut Industrial Pedagogical College, Yakutsk, Russia

Abstract. The article concerns the establishment and determination of the significance of ethno-cultural competence of students in the conditions of multicultural educational space. Based on the study of scientific literature, the article reveals the essence and content of the concepts "ethno-cultural competence" and "electronic educational resources", considers the degree of development of the problem under study, the possibilities of an open educational portal. The article describes the results of the online research among the students studying in the field of "Psychological-pedagogical education" on determining the level of formation of ethnocultural competence of students. It is revealed that at the beginning of the research the students have an insufficient level of ethnocultural competence formation. After the work on the presented set of activities through the use of electronic educational resources on the platform "Open Educational Portal of NEFU" the level of ethnocultural competence has increased. This tool has shown to be effective in introducing a dialogue of cultures, interethnic understanding and interaction, and the formation of ethnocultural competence. The Open Educational Portal of M. K. Ammosov North-Eastern Federal University is an additional means of communicative interaction between teachers and students, an information environment for the exchange of electronic resources, information, collection and simultaneous use of sources in the electronic educational environment, which provides students with the opportunity to receive quality education and develop their knowledge in various fields. #COMESYSO1120.

Keywords: Ethnocultural Competence · Electronic Educational Resource · Project Activities · Educational Portal

1 Introduction

Today, the formation of competence plays an important role in the education system, which provides students with personal and professional development. In the conditions of multicultural educational space, it is important to form students' ethno-cultural competence, which provides spiritual and moral formation and personal development of a

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 226–238, 2024.
https://doi.org/10.1007/978-3-031-53552-9_21

linguistic personality in the dialogue of cultures. Thus, the main task of education is the preservation of ethnic identity, cultural heritage and interaction, the development of ethnic consciousness and tolerant attitude towards other ethnic groups. In this regard, the formation of ethno-cultural competence, inter-ethnic attitude of students through effective pedagogical methods, ways, technologies acquires special relevance. One of such pedagogical technologies is project activities with the use of educational resources on the Internet. In this connection, we have developed an online course "Psychological and Pedagogical Fundamentals of Tutor Support" on the platform of the M. K. Ammosov North-Eastern Federal University open educational portal.

Many Russian and foreign researchers (E.P. Belinskaya, T.V. Poshtareva, O.L. Romanova, V.K. Shapovalova, E. Ericsson, G. Tadjfel, Martin Barrett, V.Y. Hotinets, A.B. Pankin, D.A. Danilov, etc.) have studied the problem of ethnocultural competence formation. They paid special attention to the fact that in order to understand and accept the values of another culture it is very important to know and assimilate one's own ethnocultural values. Thus, D.A. Danilov notes that "only a person who knows the culture and values of his people, who respects his own people, can understand and respect other peoples" [2].

Martin Barrett notes that "national identity is an extremely sophisticated complex, including various cognitive and affective as well as behavioral components. Cognitive components include such elements as historical knowledge (knowledge of historical events, historical characters, historical symbols, etc.), geographical knowledge (geography of the country, etc.), social and social psychological knowledge (knowledge of traditions and customs, knowledge of the national group, etc.)" [8].

We agree with the opinion of Y.A. Ortikov that "correct and timely formation of ethno-cultural identity in the process of globalization and enrichment of its content with various social, national and religious values will reduce the risk of disappearance of these ethno-cultural identities in the future" [9]. Accordingly, the educational system faces the task of organizing and conducting educational work for the formation of ethnic identity among students.

The interaction and connection of language, culture, mentality, and values form ethnic self-consciousness, as well as a variety of communications. D.A. Danilov notes that "ethnic self-consciousness is a set or process of evolution of a set of feelings, stereotypes (that is, emotional and psychological elements), ideas, views, knowledge (that is, rational and informational ideas) and stable ideas (that is, cognitive settings), worldview forms (that is, behavioral settings) of a person in relation to his own and other ethnic groups" [2]. The leading phenomenon of ethnic self-consciousness is ethnocultural competence.

T.V. Poshtareva defines ethno-cultural competence as "a property of a personality, expressed in the presence of a set of objective ideas and knowledge of this or that ethnic culture, realized through abilities, skills and behavior patterns, contributing to effective interethnic understanding and interaction" [6].

Based on the researchers' definitions, ethno-cultural competence can be considered as a person's assimilation of nationally specific behavioral norms, values and ideas about his/her own and another culture.

S.N. Gorshenina notes that "the essence of ethno-cultural competence lies in the existence of a number of objective ideas about ethnic culture, which allow it to adapt the specific lifestyle of specific ethnic communities, contribute to effective interethnic understanding and interaction" [1].

According to some researchers, "the introduction into the educational process of online courses of choice, having the purpose of developing the qualities of ethno-cultural communication of students will contribute to their successful adaptation in the multicultural environment of the university. The specifics of online learning also require revision of pedagogical conditions of formation of students' ethno-cultural competence, search for new organizational and content forms and methods of educational process" [7].

The aim is to determine the possibilities of electronic educational resources in the formation of students' ethno-cultural competence in the conditions of multicultural educational space.

2 Methods

We believe that in the formation of ethno-cultural competence the most effective is the project method using electronic resources. The issues of project activities in educational activity were studied by J. Dy, K.N. Polivanova, G.A. Tsukerman, N.V. Matyash, Z.S. Zhirkova. Thus, in the works of Z.S. Zhirkova we find a statement about the importance of the project method in the development of regional education system, including "in the preservation of people, its gene pool, ensuring sustainable dynamic development of Russian society - society with a high standard of living, civil, legal, professional and domestic culture" [3].

We believe that the most accurate definition to the concept of "project activity" was proposed by N.V. Matyash: "joint cognitive, creative activity, which is focused on mastering by students the methods of independent achievement of the set cognitive task, satisfaction of cognitive needs, self-realization and development of significant personal qualities within the implementation of an educational project" [5]. The main goal of project activity is to develop independence, creativity and entrepreneurial spirit of students.

The organization of project activities is productive, if you use the capabilities of digital electronic technologies.

As it is known, ethno-cultural competence can be formed by different forms: cultural events, local history conferences, ethno-pedagogical foresights, hobby groups, communities, museum visits, etc.

We present a set of regional traditional events, which provides effective formation of ethno-cultural competence of students:

- holding the Day of native language (February 13);
- exhibitions of folk craftsmen;
- cultural forum;
- International conference "Linguistic and Cultural Diversity in Cyberspace";
- Republican ecological action "Nature and We";
- Week of History of Yakut Cossacks;
- Cultural Olympiad "The Village - Cradle of Talent";

- Olonkho Performers Contest;
- development of practical assignments on project activities.

Experimental work was carried out on the basis of the Pedagogical Institute of the M.K. Ammosov North-Eastern Federal University. Students of the 2nd year in the direction of "Psychological and pedagogical education" (bachelor's degree) took part in the online study. In order to reveal the level of students' mastery of the richness and understanding of the features of native and other cultures, the ascertaining experiment was conducted in September 2022. The purpose of the online survey for diagnosing ethnocultural competence was to assess the level of formation and determine the significance of ethnocultural factors in students.

The level of knowledge and formation of ethno-cultural factors was carried out according to such indicators:

- ideas about the culture, customs and traditions of their people;
- ideas about the culture, customs and traditions of other nations;
- manifestation of ethno-cultural factors in everyday life (use of traditions in everyday life, etc.).

The online survey for students consisted of open-ended questions, to which students themselves formulated an answer, and closed-ended questions, where students chose from the suggested answers. For example, to the open-ended question "What traditional holidays, rituals, and customs of your people (Sakha) do you know?" - students named holidays that were important to them. The results were processed by the frequency of naming these or those holidays.

3 Results

It was important to check the students' knowledge and understanding of the concept of "ethno-cultural competence". To the question "What do you understand by" ethnocultural competence "?" most of the students provided detailed and clear answers.

The most vivid elements of national identity are the customs, traditions and holidays of the nation. According to the online survey, the traditional holiday Ysyakh (100%), the rite of blessing Algys, the rite of fortune-telling "Tanha ihillähine" (eavesdropping on Tanha) are highly recognized by the Sakha people. For the people of Russian nationality prevail the winter send-off holiday Maslenitsa, the folk holidays Ivan Kupala and Svyatka, the Christian holidays Christmas and Easter, the rite of Baptism; students have difficulties in questions related to the culture of small peoples of the RS (Ya) (Eveny, Evenki, Dolgan, Chukchi, Nenets): in the answers one feels uncertain, but 27% noted the national Evenki holiday Bakaldyn, the rite of hunting and fishing. To the question "From what sources did you learn about the customs, rituals of different peoples?" the answers "in the classroom" (34%), "from parents, family" (33%), "in books" (8%), "Internet, social networks" (25%) prevail.

It was important to find out what ethno-cultural awareness gives a person, as interest in the phenomenon of ethno-culture, representation and awareness of one's ethnicity plays a special role in the development of ethno-cultural competence. Thus, 83% of students have reasoned and clear answers, but 17% did not answer.

An important factor in the formation of ethno-cultural identity is knowledge of the history of the people. To the question concerning the foundation of Yakutia and the capital, students answered correctly (83%), many indicated the founder of Yakutsk, however (17%) did not give an answer. Many of them (37%) mentioned squares of Yakutsk, only 27% gave a more detailed answer and 36% did not answer the question about the history of memorable places of the city (monuments, historical buildings, streets, temples). To the question "Have you ever visited sights of Yakutia with a tour? If yes, which ones?" 70% gave the answer "yes", 30% "no". Many students indicated in their answers that they know many sights and want to visit these places.

An important element of ethno-cultural identity are folk sayings, proverbs and phrases. The question on the knowledge of proverbs and sayings, phraseological expressions caused difficulties for the students. The majority cited the most frequent, commonly known proverbs. It follows those the students' personal paremiological and phraseological fund of the language is low.

A vivid indicator of ethno-cultural competence is the interest, knowledge of the culture of another people. All the surveyed students have no idea about the specific ethno-cultural norms and rules of behavior in different cultures, which shows a low level of formation of ethno-cultural competence.

The next block implies closed questions. According to the results we can see that the students showed a clear idea of the necessity of learning intercultural interaction with representatives of other nations, dialogue of cultures, tolerant attitude and respect for another culture.

It is worth noting that in order to form ethno-cultural competence it is necessary to introduce the positive experience of interaction with other cultures, acquaintance with different nationalities, etc. from an early age. The education system plays an important role in transmitting traditional culture. And as future teachers, most students have a clear idea of what works need to be done to form ethno-cultural competence and believe that it is necessary to introduce students to the culture of other nations (Fig. 1).

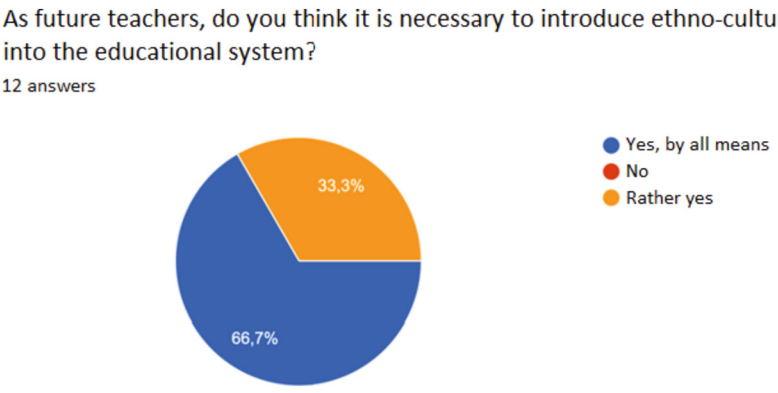


Fig. 1. Results of the closed question #1.

Earlier it was noted that the knowledge about the native culture and culture of other nations are the main indicators of ethno-cultural competence. As the results show, the students have incomplete, but sufficient knowledge (Fig. 2).

As we see in Fig. 3, most students adhere to the traditions and customs of native culture in everyday life (Fig. 3).

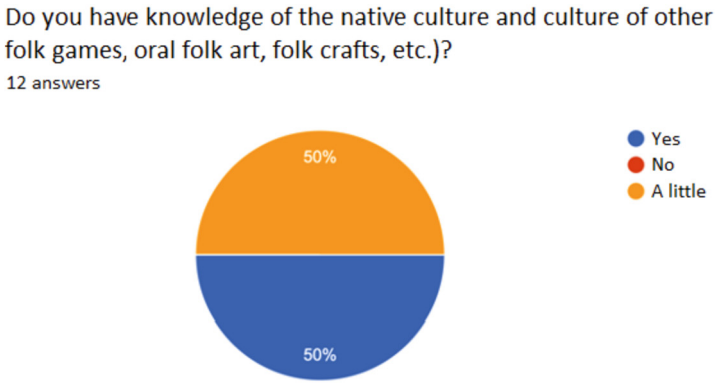


Fig. 2. The results of the closed question #2.

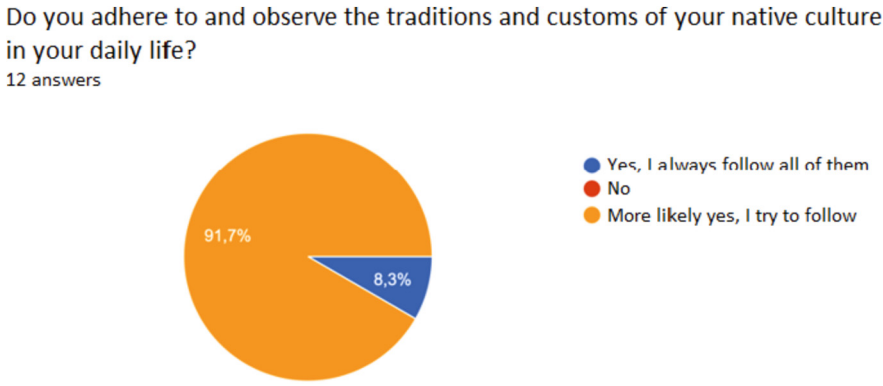


Fig. 3. The results of the closed question #3.

Also, national traditions and customs are observed in the family (Fig. 4).

The majority of students believe that first of all nationality is determined by culture, traditions and everyday life, then by mother tongue (Fig. 5).

Students try to attend national holidays and events (Fig. 6).

In general, students have a desire to learn about the traditions and culture of a foreign nation, but 16.7% do not aspire to know the traditions and culture of another nation (Fig. 7).

In general, many students have an idea of culture, traditions and customs of their own and Russian people, but have difficulties with representatives of minorities. According

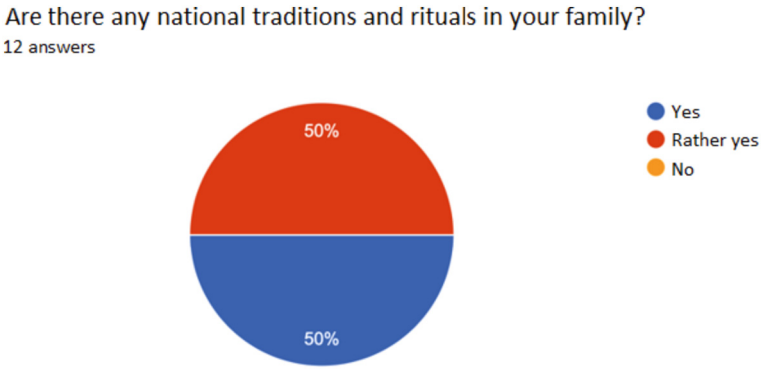


Fig. 4. Results of the closed question #4.

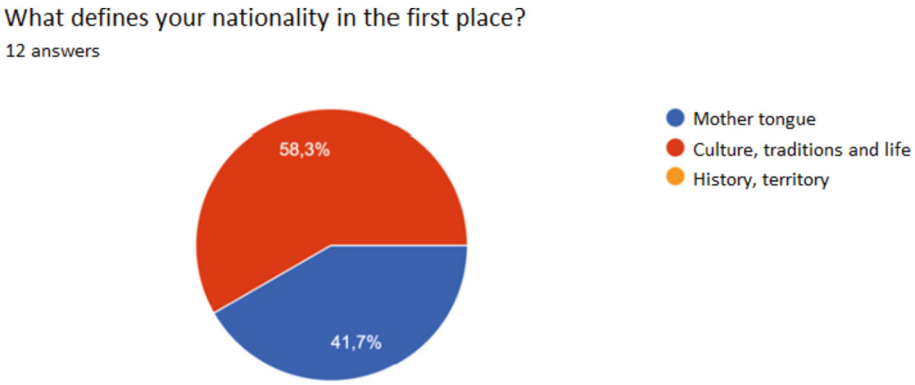


Fig. 5. Results of the closed question #5.

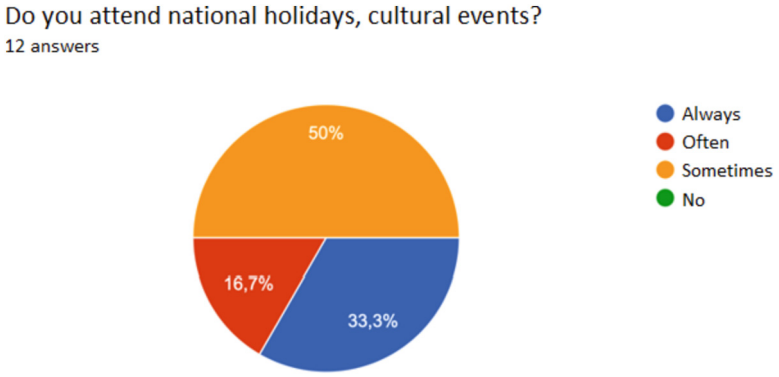


Fig. 6. The results of the closed question #6.

to the results of the online study, we can understand that many are ready for cultural

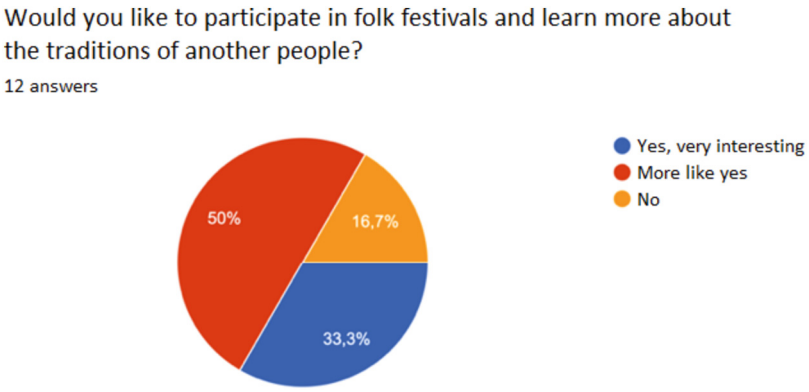


Fig. 7. The results of the closed question #7.

interaction, intercultural communication. As we can see, students adhere to and keep traditional rules, first of all, it concerns festive culture and rituals. Students believe that the culture and traditions of the people play a determining role in the self-identification of peoples than the language. However, they do not have a clear and complete knowledge of history, memorable places, as well as about the peculiarities of language units.

One of the important means of forming ethno-cultural competence is purposeful modeling of the dialogue of cultures in the classroom. It is important to note that the identity and uniqueness of the culture of the people is manifested by comparing and identifying the commonalities of the native culture and the culture of the other people. The dialogue of cultures is carried out through the performance of various means and methods. Thus, the priority is the method of project activities with the use of various electronic educational resources.

The educational environment of higher education plays an important role in the formation of ethno-cultural competence and intercultural interaction. Educational institutions educate people from different cultures, which provides an opportunity to exchange experiences and cultural interaction. In the process of education students can learn a lot about other cultures and traditions, which helps to form ethno-cultural competence, tolerant attitude towards the ideas of other nationalities. Thus, an important factor in the formation of ethno-cultural competence is a variety of educational programs and online courses. Online courses can contain different cultural traditions that teach students to understand the values inherent in different cultures, to interact in everyday, academic, and later professional contexts.

4 Discussion

The conducted research on the formation of ethno-cultural competence allowed us to determine the goals and content of the online course “Psycho-pedagogical foundations of tutor support”. The online course is implemented in the process of extracurricular work with students on the platform of the NEFU Open Educational Portal.

The NEFU Open Educational Portal is an educational platform of online courses offered by NEFU for a wide range of students from all over the world. The educational portal is an additional means of communicative interaction between teachers and students, an information environment for the exchange of electronic resources, information, collection and simultaneous use of sources in the electronic educational environment. The open educational portal for students provides an opportunity to receive a quality education and to develop their knowledge in various fields. With its help students can study at any convenient time and from any place in the world. The main features of the NEFU Open Educational Portal include: taking courses to improve qualifications, professional retraining; the possibility of subject preparation for the unified state exams; attracting the attention of young people to the sphere of high technology, helping students of general education schools and students of secondary vocational education to master modern information and communication technologies; participation in various Olympiads.

The portal contains many educational resources, such as video lectures, online courses, tests and assignments. Thus, the method of projects using electronic resources is an effective way to stimulate and develop communicative competences, the skill of searching and processing information from various sources of the Internet. In the process of working in the educational portal, the teacher acts as a consultant, a tutor of students' project activities.

The use of educational resources in the learning process based on the project method contributes to:

- formation of communicative skills and abilities to carry out information work with computer equipment, various software;
- development of critical, systematic and analytical thinking;
- active use of different electronic technologies in the learning process;
- ability to analyze and systematize the material in the process of creating projects.

The online course is focused on the development of project activities. In the process of training students will get necessary knowledge and practical skills for effective implementation of projects, development of team skills, as well as for successful solution of problem situations during project activities.

The online course includes the following topics:

- What is a project: an introduction to project activities;
- Tutor approach and its application to project activities in the education system;
- Development and implementation of projects;
- Working with project results and their evaluation.

In the process of training students will have an opportunity to receive feedback on their assignments and works from the teacher, which will help them even more confident and qualitatively perform tasks related to the tutor's support of projects.

The process of organizing project activities takes place under certain pedagogical conditions: increasing students' motivation and interest in creating projects, systematic and activity-based approach, availability of creativity and originality of projects, performing special cultural tasks, observing a clear algorithm of project stages. The material of the online course contains an ethno-cultural component and is aimed at solving the

following tasks: formation of students' ethno-cultural competence; development of communicative and creative competence; formation of positive ethnic identity of students, etc.

Thus, by performing project activity tasks that contain ethno-cultural components, students are involved in the process of consciousness with the traditional values of Russian and native culture.

Directions for the implementation of project activities may include: presentation of the history of traditions and customs, cuisine, national costume of their own and other nations; familiarization with famous names and works of writers, bearers of epic traditions of the people, prominent figures of their and other nations.

One of the effective means of forming students' ethno-cultural competence in the classroom is systematic work with language units with a national-cultural component of meaning, in particular with phraseological units. The role of the national-cultural component in the formation of the semantics of phraseological units is important as it reflects the specificity of national mentality, values, ideas and stereotypes of this or that nation allows to perceive and transmit information, which develops ethnocultural competence.

As an example, we present the materials used in the online course.

"Project work: description of concepts: light - darkness, good - evil, beauty, truth - verity, happiness, soul, friend (or choose another word).

Project plan:

1. My associations with the word (the word in my mind)
2. Dictionary portrait of the word (word in dictionaries: explanatory, etymological (source of origin), phraseological (with examples, with illustrations, give an analogue from the Yakut language), synonyms, antonyms).
3. Answer the question: do the given phraseological expressions reflect the peculiarities of Russian culture? Are the phraseological expressions connected with the history and culture of the people?
4. Textual portrait of the word (aphorisms with the word, lines from poems, works).

Project work.

Group 1: Phraseological expressions whose origin is connected with the customs and rituals of our ancestors.

Group 2: Phraseological expressions originated in one or another professional environment or came into the literary language from jargon.

Group 3: Phraseological expressions of ancient origin (from mythology, literature, history).

Group 4: Phraseological expressions of biblical origin.

Group 5: Phraseological expressions of folklore origin.

Group 6: Phraseological expressions whose origin is connected with the history of our country.

Plan of research of phraseology:

1. Name the phraseology chosen for the study.
2. Identify its source.
3. Explain the meaning of the phraseology.

4. Tell the history of the origin of the phraseology.
5. Give an analogue of the phraseological phrase from the Yakut language, explain the meaning (if there is one).
6. Find a synonym, an antonym (if there is one).
7. Gather and draw up information about the phraseology (illustrate)" [4].

Project work "Creating an online dictionary". There are phraseological expressions that reflect the history of the people, religion, national proper names. Each language reflects its worldview, reflects the life and activities of each nation.

Let us make sure of this and on the example of the native language. With the help of the Yakut-Russian phraseological dictionary, determine the origin of these phraseological expressions, and name the fact of Yakut history (custom) that formed the basis of the phraseological expression. Are there similar facts (custom/belief) in Russian culture?

Ayyy sire ahaa ҕas, k ү n sire k ү nday (lit. Country of deities is open, sunflower country is capacious).

Kirdeh kihi (lit. Dirty man).

Sakha a ҕabyyta (lit. Yakut pop).

Bayanaidaah bulchut (lit. Having Bayanai hunter).

Ayy kihite (lit. God's man).

Aiyy siriten ara ҕysta (lit. he parted from the god world).

Project work (creative). Below are phraseological expressions in Russian and Yakut. What human qualities do the phraseological phrases connected with pets usually denote? How can we determine the meaning of phraseological expressions (with animals: chickens, dog, horse, mouse, fox). What is the cultural significance of these words? Use phraseological dictionaries, illustrate Table 1.

Table 1. Phraseological expressions in the Russian and Yakut languages.

Money's tight	Böröm bahá, bá ҕam bahá (lit. my wolf's luck, my frog's saber)
Let's get back to our sheep	Kutuyah iinin (khoronun) keketer (lit. Digs a mouse hole)
Ate the dog	Kermes sahyl meldirere (lit. Denying the foxy sivodushka)
Don't drive the horses away	Kyrsa sahyl o ҕoto (lit. Fox cub, baby fox)
Dark horse	Bader meii (lit. Lynx brain)
	Tii ҥmeii (lit. Squirrel memory)

At the control stage of the study, we repeated the online survey. According to the results, we can see that the dynamics of the level of knowledge and possession of ethnocultural information, and, consequently, the formation of elements of ethno-cultural competence after the presented sets of activities using electronic resources is observed.

A comparative characteristic of the results of the ascertaining and controlling stages of the experiment is presented in Table 2.

Table 2. Comparative results of closed questions.

	Question 1	Question 2	Question 6	Question 7
Contesting stage	Yes (66.7%) Rather yes (33.3%)	Yes (50%) A little (50%)	Always (33.3%) Often (16.7%) Sometimes (50%)	Yes, very interesting (33.3%) More likely yes (50%) No (16.7%)
Control stage	Yes (83%) Rather yes (17%)	Yes (82%) A little (18%)	Always (46%) Often (36%) Sometimes (18%)	Yes, very interesting (59%) More likely yes (33%) No (8%)

5 Conclusion

Our research has shown that the formation of ethno-cultural competence contributes to the students to the dialogue of cultures, effective inter-ethnic understanding and interaction, the formation of ethno-cultural competence. It also develops a sense of pride in their people, ethnic identity, patriotism. The results of the experiment proved that the use of electronic educational resources, in particular the online course "Psycho-pedagogical foundations of tutoring", containing ethno-cultural material, is an effective method of forming ethno-cultural competence. Mastered knowledge and skills to work with electronic resources, the process of implementation of project activities increases the level of professional training of students, forms the ability to organize the learning process in a digital educational environment.

References

1. Gorshenina, S.N.: Technology of pedagogical support for the formation of ethnocultural competence of schoolchildren in the conditions of multicultural educational space. TSU Sci. Vector **2**, 30–33 (2015)
2. Danilov, D.A., Ivanov, A.V., Chabyev, I.P.: Development of Ethnic Self-Consciousness of Adolescents in the Pedagogical Process: on the Material of the Yakut School. YSU Press, Yakutsk, Russia (2003)
3. Zhirkova, Z.S.: Management Strategy for the Development of the Regional Education System by Means of Innovative Design. Academia, Moscow (2011)
4. Ivanova, S.S.: Linguocultural tasks as a means of forming ethno-cultural values in bilingual students. In: Proceedings of the All-Russian Scientific Conference "Ethno-pedagogy as a factor in the preservation of Russian identity", pp. 45–49. ICITO, Kirov, Russia (2022)
5. Matyash, N.V.: Project Activity of School Children. High School, Moscow, Russia (2000)
6. Poshtareva, T.V.: Formation of ethno-cultural competence. Pedagogics **3**, 35–42 (2005)
7. Arsaliev, S.M.-Kh.: The development of ethno-cultural competence of university students during COVID19 pandemic in Russia. In: Proceedings of the International Conference on Social Science and Education Research 2020, ICSSER, pp. 46–50. Atlantis Press, Netherlands (2020)

8. Barrett, M.: Development of National, Ethnolinguistic and Religious Identities in Children and Adolescents. Institute of Psychology, Russian Academy of Sciences (IPRAS), Moscow, Russia (2001)
9. Ortikov, Y.A.: Ethno cultural identity and its formation. Asian J. Multidimension. Res. **10**(4), 516–522 (2021)

A Model of Analysis of Daily ECG Monitoring Data for Detection of Arrhythmias of the "Bigeminy" Type

D. V. Lakomov[1], Vladimir V. Alekseev[1(✉)], O. H. Al Hamami[1], and O. V. Fomina[2]

[1] Tambov State Technical University, Tambov 392000, Russia
vvalex1961@mail.ru
[2] Tambov State University Named After G. R. Derzhavin, Tambov 392036, Russia

Abstract. The article presents the results of a study on the creation of a neural network model of a big data processing system obtained from daily ECG monitoring devices (Holter sensor) to search for characteristic points on the ECG graph. A method of overcoming the noise level of the initial data found in the course of the study is described. Using the example of the search for arrhythmias of the "bigeminia" type, the high efficiency of the model is shown, combining the use of a data preprocessing algorithm to eliminate noise and the use of a fully connected four-layer neural network. The presented results of a computational experiment to test the effectiveness of the model in organizing decision support by a cardiologist about the absence or presence of bigeminia are consistent with previously known theoretical provisions and experimental results in this area. As a result, the direction of further research for the creation of a decision support system by a cardiologist in the analysis of these technical means has been determined. #COMESYSO1120.

Keywords: Bigeminy · Big Data Analysis · ECG · Holter

1 Introduction

At present, the development of decision support systems (DSS) for physicians is particular importance. When artificial intelligence has entered the field of medicine in recent years, machine learning approaches have made progress in helping medical professionals to optimize personalized treatment in a given situation, in particular the electrocardiography and the interpretation of the result. Artificial intelligence methodologies are increasingly being introduced into all aspects of patient monitoring and are paving the way to minimally invasive or non-invasive treatments and diagnostics. The application of artificial intelligence in the analysis of images and big data processing, obtained in performing of medical functional diagnostics of a person, allows to increase the accuracy and speed of diagnosis, as well as provides the ability to predict the course and the development of the disease or vice versa, recovery of the body. In particular, the application of artificial intelligence turns the ECG into a ubiquitous, non-invasive cardiac test, that

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 239–244, 2024.
https://doi.org/10.1007/978-3-031-53552-9_22

is integrated into the workflows of practice, a screening tool, and a predictor of heart and non-cardiac disease in asymptomatic people. The application of artificial intelligence to the Holter monitoring protocol makes it possible to diagnose conditions, that were not previously identified by an electrocardiogram, or to do it with greater productivity. As well as to predict the development of diseases or complications based on the developed neural network and created knowledge base [1, 2].

A few words about the medical aspect of this problem. Bigeminia is a type of allorhythmy, i.e. regular repetition of extrasystoles [1].

Ventricular extrasystole is caused by a pathological excitation wave emanating from the ventricular conductive system.

Ventricular bigeminy has the following electrocardiographic features: premature appearance of a wide and deformed complex QRS (PQRST of ECG teeth); compensatory pause, following the complex, in the form of a straight line; lack of tooth P; direction of the ST segment to the side, opposite to the QRS complex [2].

2 Methods

The basic data for solving the problem of detecting "bigeminy" arrhythmia were obtained from the Holter monitoring. The database consists of numbered folders with daily ECG data of several dozen people. In which episodes of bigeminy occur in a significant number. Each folder was represented by two files:

– file ECG. Txt.
– file dataset.xlsx.

The ECG.txt file contains the 7-lead ECG graph values were written in columns separated by a space in the following order: I, II, V5, III, aVL, aVR, aVF. The lead values are recorded in the same order for all patients with a frequency of 266 Hz.

The dataset.xlsx file contains a table with information on the data decryption from the ECG file, which contains the following information:

- time of the fixed complex in milliseconds from the beginning of the day;
- time of the fixed complex in milliseconds from the beginning of recording;
- line number in the ECG file with values corresponding to the fixed complex;
- type of complex found: S-supraventricular; V - ventricular;
- a brief designation of the fixed arrhythmia;
- the name of the arrhythmia.

3 Results

As part of the study, we analyzed the available data. The regularity of the length of the complexes was revealed using the available values. The average value of the complex was 460–600 ms, which corresponds to the signs of tachycardia [1].

During the analysis, the noise of the source data was detected. To eliminate it at the stages of preprocessing and data filtering, we applied a two-step method:

1. At the first stage, emissions for the 1st and 99th percentile were discarded (see Fig. 1). All data that is above and below these lines are assigned a zero value. Between 15:00 and 17:00 from the start of shooting, emissions and the end of recording are visible. Probably at this time, the patient was asleep and could not control the position of the electrodes. As part of this stage, the values labeled "x and z artifacts" were excluded from the primary data sample for rows with the "bigeminy" characteristic (see Fig. 2).
2. At the second stage of pre-processing for the complexes, the values were converted into positive in the form of "1" and negative in the form of "−1" in the 5th and 95th percentiles, the remaining values were equated to "0" (see Fig. 3 and Fig. 4).The second stage is the determination of the tooth R on the ECG, because indicators take the highest or the lowest values in this spike. In our study, the 5th percentile is responsible for the lowest values, and the 95th percentile is responsible for the highest. After finding the tooth R, we can proceed to determine the remaining components of the PQRST complex, which will be necessary to build a complete ECG model [3].

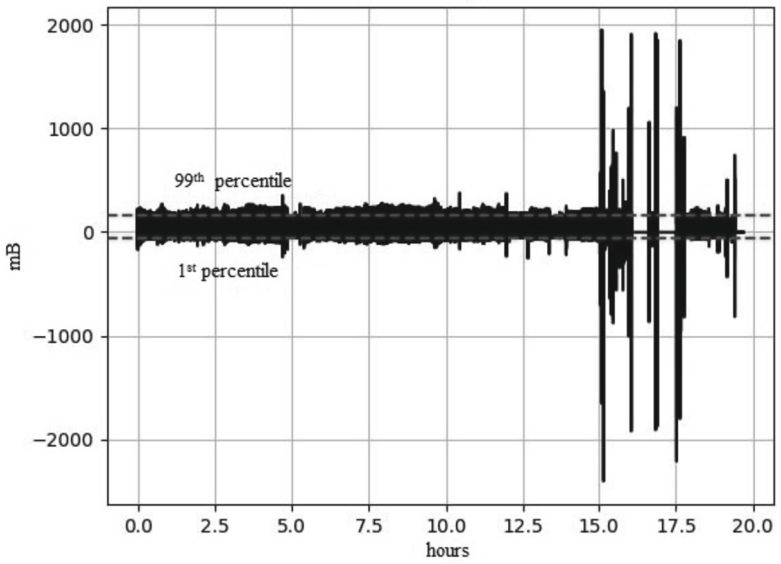


Fig. 1. Primary data analysis

Graph in Fig. 4 simulates the ECG and shows already fully filtered data for the operation of the neural network.

The pre-processing method mentioned above made it possible to filter out most of the noises and reduce their impact on the main data sample.

The main element of the program module comes next – a fully connected neural network with 4 hidden layers [4]. Statistical data analysis has become an additional element for identifying lines with parameters inherent in bigeminy. Pre-processing of data made it possible to increase the value of the accuracy of identifying parameters characteristic of bigeminy to 89%. With further interpretation of the available data and

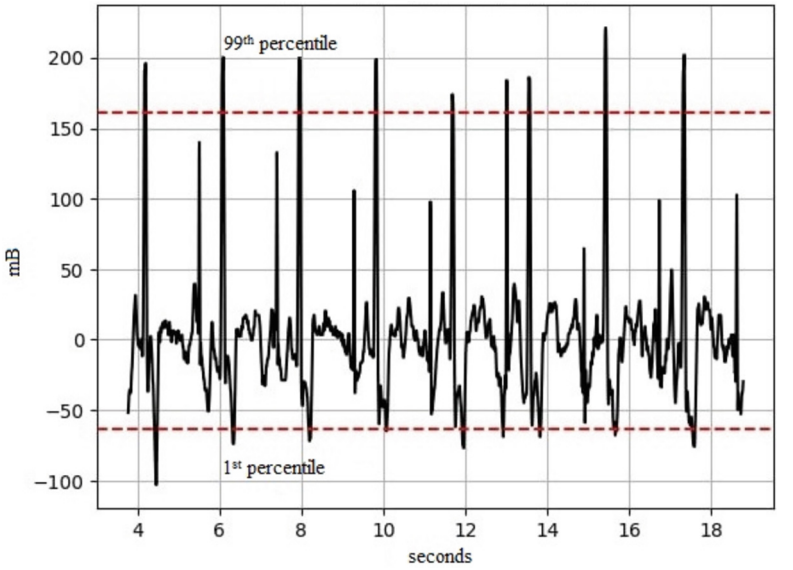


Fig. 2. Adjusted primary data sampling graph

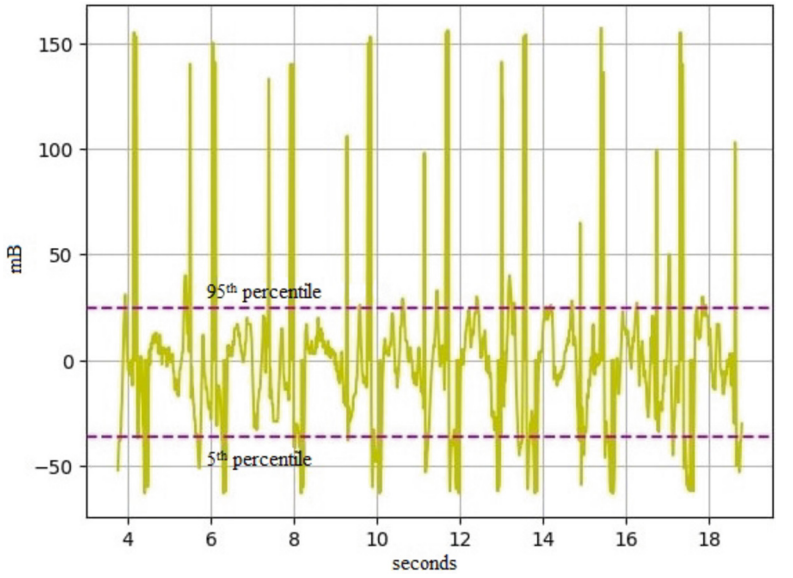


Fig. 3. Preprocessed ECG

calculation of the length of the complex with the arrhythmia parameter, it was revealed that with the emerging bigeminy, the length of the complex averaged 900–1340 ms,

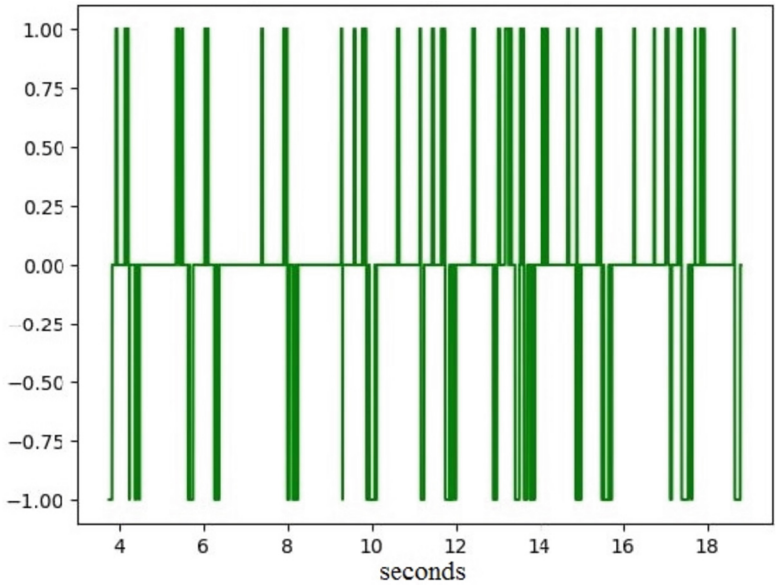


Fig. 4. Data for model training

which corresponds to the values of the full compensatory pause, which is one of the signs of bigeminy.

4 Conclusion

In the course of the study, the team of authors proposed an approach for processing data obtained from the Holter monitoring, which includes an algorithm for pre-processing data to eliminate noise and the use of a fully connected 4-layer neural network that searches for arrhythmias of the "bigeminy" type.

As a further development of the developed model, it is advisable to use the LAMA neural network additionally to reduce time costs and increase the productivity of the system.

Based on the further experiments, it is advisable to develop and train a logical-linguistic model that would allow to maximize the percentage of finding arrhythmias of the "bigeminy" type without increasing computational costs.

Further development of the model may also be associated with the creation of a fully functioning decision support system to help the cardiologist to focus on some parameters that are important According to the system. It should be noted separately that the model should not replace the doctor, but work as a recommendation system.

References

1. Rudenko, M.Y.: Cardiometry. Fundamentals of Theory and Practice. ICM, Taganrog, Moscow (2020)
2. National Russian recommendations on the use of the Holter monitoring technique in clinical practice. Russ. J. Cardiol. **2**(106), 6–71 (2014)
3. Polyakov, D.V., Eliseev, A.I., Andreev, D.S., Selivanov, A.Yu.: Machine learning of models of fuzzy classification. In: Engineering Physics, No. 7, pp. 19–35, Nauchtehlitizdat, Moscow (2021)
4. Alekseev, V., Eliseev, A., Kolegov, K., Aminova, F.: Neural network algorithms to control dynamic objects. In: 2019 1st International Conference on Control Systems, Mathematical Modelling, Automation and Energy Efficiency (SUMMA), p. 712. Lipetsk State Technical University, Lipetsk (2019)

Development of a Data Mart for Business Analysis of the University's Economic Performance

Sergei N. Karabtsev(✉), Roman M. Kotov, Ivan P. Davzit, Evgeny S. Gurov, and Andrey L. Chebotarev

Kemerovo State University, Kemerovo, Russia
skarab@kemsu.ru

Abstract. In higher education there exist numerous strategic tasks impossible or too labor intensive to be solved by traditional methods, with one being financial planning, which allows for all planned activities during the academic year. Simultaneously, a university's total budget is not static and can be adjusted several times a year due to changes in the student body. Furthermore, a decrease in the number of students due to academic failure or an incorrectly defined enrollment plan for specific training programs can result in this program becoming economically unfeasible due to the costs outweighing the income contingent on the size of the student body. The objective of this work is using data management to improve monitoring and forecasting the economic performance at Kemerovo State University. Universities accumulate a vast amount of heterogeneous data, the analysis of which can ensure decision-making based on data, not on intuition. Analysis of large amounts of data is impossible without the use of modern products and technologies pertaining to Business Intelligence. The paper sets the task of creating a Decision Support System (DSS) related to filling the university's budget through enrollment or expulsion of students and the economic evaluation of specific training programs. Research methods involve creating a DSS with a description of the main results on each stage alongside statistical data analysis. Introducing an information system at the university made it possible to predict student retention and potential budget losses, evaluate economic indicators in the implementation of training programs, and make decisions on their adjustment. #COMESYSO1120.

Keywords: Decision Support · Analysis of Economic Indicators · Business Intelligence · Data Mart · Dashboard · ETL

1 Introduction

Higher education institutions face a great number of challenges, caused by the current social and economic conditions, state education policy, and the increasing competition between universities, leading to the necessity to position oneself at the top of various authoritative ratings lists. Other major tasks may include attracting highly qualified specialists, improving academic performance and satisfying the educational demands

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 245–258, 2024.
https://doi.org/10.1007/978-3-031-53552-9_23

of students, financial planning and increasing the share of extrabudgetary resources in the total budget, promoting and commercializing cutting edge research. It is also of utmost importance to attract applicants in the face of internationalization of education and the remote nature of the way students interact with the university during admission campaigns. Traditional approaches to solving such tasks may either be inapplicable due to their high labor intensity or entirely inadequate. Modern approaches that allow tackling the above-mentioned tasks rely on IT, methods and technologies for analyzing the data accumulated at the university for subsequent decision-making that is based on data, not intuition.

Over the course of their activity, every day higher education institutions create, collect, accumulate and store a large amount of data on a student's academic performance, the faculty's educational and extracurricular workload, their research activity, publications, the university's financial operations and budgetary planning, and so on. At best, the vast majority of this data is stored in a digital form in disjointed information-gathering-and-tracking systems and databases, and at worst – on physical media, which is problematic for data analysis algorithms and subsequent decision-making, since autonomous databases and data sources are unable to give consolidated information clearly enough. Furthermore, the data collected may be incomplete, contain inaccuracies and discrepancies with the real business processes at the education institution. Before data becomes usable for the purposes of correct analysis, it is necessary to perform a number of organizational and technical activities, beginning with developing a university digitalization strategy, accounting for the aims and objects of the analysis, detailed university performance indicators, and culminating in IT solutions – creating data marts or data warehouses and using Business Intelligence (BI) tools for analysis [1–3].

A well-known type of BI tools is OLAP – On-Line Analytical Processing, which is used in descriptive analytics [4, 5]. Descriptive analytics focuses on historical data, aiming at finding answers to the questions pertaining to what happened, why it happened and what is happening at the moment. Having become the foundation for making educated decisions, analytical data processing has led to appreciating the value of this process and changed the attitude towards data collection and storage. While in the past accumulated data was treated as a by-product of business activity, now this data has become a mandatory asset in an organization, playing key role in managing the company as a whole and its business processes [6, 7].

The aim of this study is improving the process of monitoring and forecasting economic indicators at Kemerovo State University (KemSU) on the basis of data management. The paper presents a Decision Support System (DSS), connected with filling the budget at the expense of enrollment or expulsion of applicants, alongside economic assessment of implementation of specific training programs under current circumstances.

The system is part of the information-situational center at KemSU that is being created now using BI technologies [7]. The objectives of the study are developing a Decision Support System for monitoring economic indicators at the university and its implementation in the daily workflow of the employees involved in decision-making. The system must correspond with the digitalization strategy adopted at the university and present numerical or graphical data in response to the following:

– the number of state-funded places and the budget allocated to the university for training program implementation;
– the number of contracted places and the expected amount of extrabudgetary income;
– academic performance and the predicted number of students to be expelled for academic failure;
– projected amount of shortfall of budget funds due to expulsion;
– the amount of funds spent implementing the training program, including the faculty's wages, as well as utility and other expenses;
– analysis of the economic feasibility of implementing training programs.

All information should be categorized by institutes, departments and training programs. It should be possible to filter information by year, form and type of education. In addition, the information system should allow different level decision-makers to use it: the rector and administration, institute directors, department heads. The decision maker must be able to access data through a thin client – a web browser.

2 Research Methodology

Developing a Decision Support System requires carrying out preliminary investigation of the university in correspondence with the aims and objectives of the study. As a result of said investigation, the level of IT infrastructure, processes, types and sources of data for analysis, and the parties interested are identified. The classic approach to realizing such a DSS system is shown in Fig. 1 and consists of three stages which reveal the source of the data required, extraction and preliminary processing before placing in a unified data storage, building a multidimensional data model and creating analytical reports representation [2, 5].

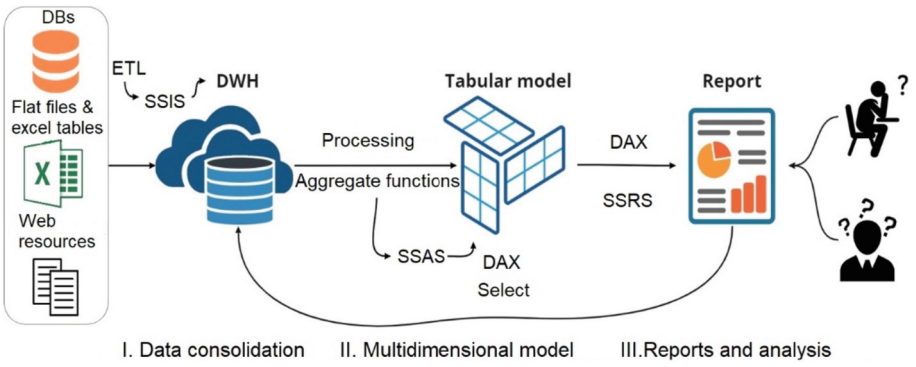


Fig. 1. Conceptual scheme of realization business-analysis in an organization.

As a DSS development tool, the study uses a BI platform by Microsoft – one of the three leaders amongst analytical and BI platforms according to Gartner. The platform consists of the following components: SQL Server 2019, Analysis Services, Integration

Services, Data Tools and Power BI Report Server [8, 9]. The necessary licenses were purchased by the university under the Microsoft Software Assurance program.

The SQL Server RDBMS is used to manage data in a database implemented as a domain-specific data mart. The structure of the data mart is designed in such a way that the information system has the ability to provide answers to a wide range of questions identified above in Sect. 1. Visual Studio 2019 with Data Tools installed is the main development tool for a BI project, consisting of data migration solutions and a multidimensional model building solution. Using the packages (SSIS packages) created by the authors of the article, the Integration Services component cleans, transforms, and transfers data from online data sources to the data mart. Besides SSIS packages, the authors of the paper created their own mechanisms implemented in Python3 using Pandas, BeautifulSoup, Pyodbc and other libraries to extract data from web resources or xml-files. Analysis Services is required to design a multidimensional model, the information from which is displayed in textual or graphical form on dashboards designed and implemented by Power BI tools. Dashboards projects are published to Power BI Report Server directories and are available to users through an internet browser and the https protocol.

2.1 Data Consolidation

Consolidation begins with searching for the available data sources, suitable for analysis. KemSU has a large number of information systems, supporting the education process, keeping track of the faculty's and students' personal files, performing financial activities, fulfilling the requirements of regulatory authorities for licensing, accreditation, scientific reporting. As a result, around 25 data sources were identified at the university. However, some of them are not automated systems, or contain aggregated data only, for example, total utility costs. In addition, most information systems do not have a developed API for interacting with external systems, making extracting data possible only at the level of direct access to the DBMS that ensure their operation (for example, using SSIS packages) or using specially created Python scripts that parse web application pages.

The following steps were taken to ensure the completeness and quality of the data. Firstly, among the 25 sources found, only those sources were selected that contain up-to-date data, and the information systems that ensure their operation are used regularly by a large number of university employees. For example, information about a student's progress can be obtained from two university information systems: one is filled in by teachers (for each training session or control event), and the second is filled in by employees of the directorates of the institute, based on these data, orders are formed to transfer to the next course or expel the student. Preference was given to the second system, since data is received in it aggregated from the first system and undergoes additional control through the actions of directorate employees. Secondly, those data sources that cannot be replaced by alternative ones were manually checked and corrected by employees. An example is an information system that processes information about teachers. The database of this system was carelessly filled with data on the scientific degree, scientific title and position of the teacher, and this could lead to incorrect billing of the cost of his work. Of course, up-to-date data is contained in the databases of the personnel department of employees, however, access to such databases is very limited, and the database

itself does not contain a link to the teaching load of the teacher in each direction of study. Thirdly, the missing data were created specifically with the help of developed information systems.

As the task was to analyze various financial costs associated with implementing the training program, it became necessary to develop an information system for Chief Financial Officer, which allows such costs to be taken into account by expenditure items, faculty staff, training programs or specific events (Fig. 2).

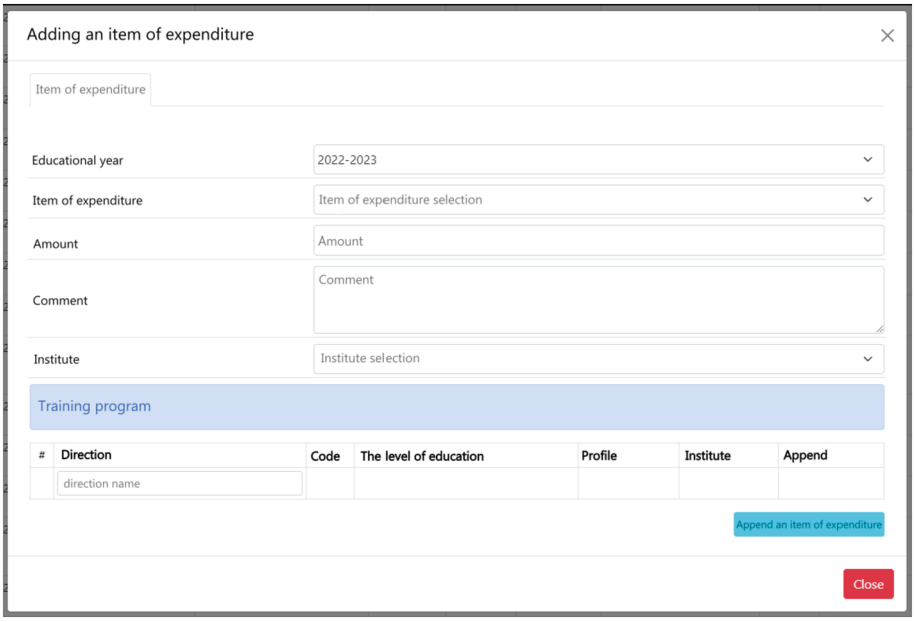


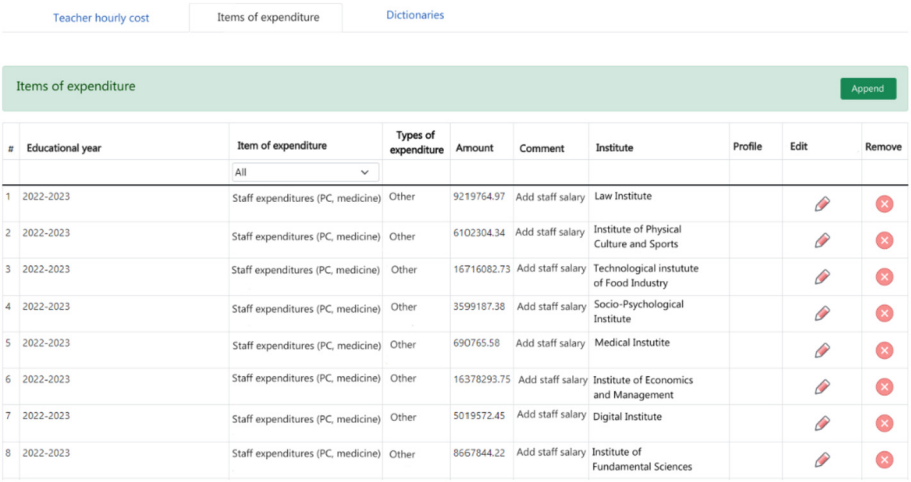
Fig. 2. Form for adding expenditure items.

Figure 2 shows a page of information system that allows one to add expenditure items, such as the cost of a medical examination or a professor's further training program, a professor's travel expenses during the admission campaign, utility costs for repairs and maintenance of premises, expenses for providing a material base for the implementation of the training program, etc.

Figure 3 shows the results of the distribution of other costs caused by carrying out the training program, correlated with the institute, expenses type, academic year (as the information is confidential, specific indicators in the figure are distorted and serve an illustrative purpose). Data is entered into the system monthly by Chief Financial Officer.

Analysis of the available data sources and technical capacity to access them revealed the following sources and data extraction approaches (Table 1):

At the next step, after determining the data sources, it is necessary to design the data warehouse structure that can ensure the implementation of the functions of the information system being developed. In this study, instead of a data warehouse we are using a data mart, the key difference of which is its simplicity in implementation and



#	Educational year	Item of expenditure	Types of expenditure	Amount	Comment	Institute	Profile	Edit	Remove
		All							
1	2022-2023	Staff expenditures (PC, medicine)	Other	9219764.97	Add staff salary	Law Institute			
2	2022-2023	Staff expenditures (PC, medicine)	Other	6102304.34	Add staff salary	Institute of Physical Culture and Sports			
3	2022-2023	Staff expenditures (PC, medicine)	Other	16716082.73	Add staff salary	Technological instutute of Food Industry			
4	2022-2023	Staff expenditures (PC, medicine)	Other	3599187.38	Add staff salary	Socio-Psychological Institute			
5	2022-2023	Staff expenditures (PC, medicine)	Other	690765.58	Add staff salary	Medical Instutite			
6	2022-2023	Staff expenditures (PC, medicine)	Other	16378293.75	Add staff salary	Institute of Economics and Management			
7	2022-2023	Staff expenditures (PC, medicine)	Other	5019572.45	Add staff salary	Digital Institute			
8	2022-2023	Staff expenditures (PC, medicine)	Other	8667844.22	Add staff salary	Institute of Fundamental Sciences			

Fig. 3. Financial expenditure items.

Table 1. Data sources and data extraction approaches.

#	Data source	Retrieved data (in general)	Data extraction approaches
1	Information system for calculating teacher workload (official website of the developer https://www.mmis.ru/)	Surname, name, patronymic of the lecturer, institute, department, volume of educational and extracurricular workload, plan of publication activity	Using direct database access MS SQL Server
2	Accounting information system (official website of the developer https://v8.1c.ru/stateacc/)	Surname, name, patronymic of the teacher, position, academic degree and title, official salary and other payments	Parsing xml-file
3	Electronic information educational environment KemSU (own development)	Surname, name, patronymic of the lecturer, lecturer's ORCID, SPIN-code, AuthorID and ResearcherID	Using direct database access Oracle
4	Information system Applicant (own development)	Competitive groups, admission targets by years, tuition fees by years	Using direct database access Oracle
5	Financial manager information system (own development)	Financial expenses	Using direct database access Oracle

the volume of data it can store. One of the most powerful information flows in the data

mart is the input flow associated with data transfer from sources [10]. Information is processed when transferred to the data mart: data is verified, cleaned, sorted, grouped and enriched by adding new attributes. This process is known as ETL – Extraction, Transformation and Loading.

In this study, ETL is carried out with the help of SQL Server Integration Service (SSIS), where packages, containing control flows and data flows, are created (Fig. 4). Control flow means transferring data in a relational mart table, and it can be done in parallel or sequentially with other flows. Data flow, in turn, refines the control flow with a specific sequence of actions: data source extraction, a series of transformations and loading into a data sink.

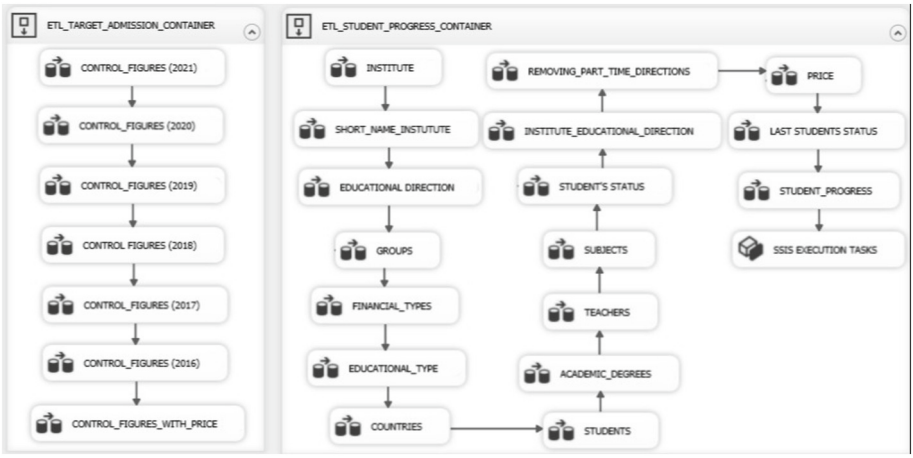


Fig. 4. SSIS Data control flows.

Data sink can be a target data mart table or some intermediary table or a temporary file which can later act as a source of transformed data. ETL is one of the most labor-intensive stages of data mart creation, accounting for up to 60% of development time.

2.2 Multidimensional Model

At the data analysis stage (Fig. 1), it becomes necessary to build dependencies be-tween the analyzed value and a set of parameters, also known as measurements. The plurality of dimensions implies presenting data as a multidimensional model, which can be visually represented as an n-dimensional cube. Measurements are plotted along the axes of the cube, and at the intersection of the measurement axes there is data that quantitatively characterizes the analyzed quantities (facts) – measures. Thus, the model for data analysis contains two main entities – measures and dimensions. The most common methodology for designing a data mart in a relational DBMS and storing multidimensional data in it is the "star" or "snowflake" schema, which is quite fully described in [10, 11].

The database schema, alongside the data, underlies the multidimensional data model [10, 12]. In this study, for simplicity and speed of development, a simplified version

thereof was used, namely, the Tabular Model. This model is reinforced by a few dozens of calculated columns and measures. An example of some calculated columns and measures is shown in Table 2.

Table 2. List of measures and calculated columns.

Measure/calculated column name	DAX-expression
Measures	
Budget	(SUMX(FILTER(STUDENT, [Form of financing] = "budget"), [money])*2)/1000000
Number of unsatisfactory grades	COUNTAX (FILTER (STUDENT_PROGRESS, [mark] = "unsatisfactory grades"), [mark])
Reduced student body	COUNTAX(FILTER('STATUSES', [Form of study] = "Full-time"), [id_student]) + (COUNTAX(FILTER('STATUSES', [Form of study] = "Full-time part-time"), [id_student])*0.25) + (COUNTAX(FILTER('STATUSES', [Form of study] = "Part-time"), [id_student])*0.1)
Current income	[Current income (budget)] + [Current income (contract)]
Current income (budget)	SUMX(FILTER('STATUSES', AND (AND(NOT([Status] = "On academic leave"), STATUSES[date] = TODAY()), [Form of financing] = "budgetary")),[money])
Marginality	([Budget and contract]-[Expenses]) / [Budget and contract] * 100
Calculated columns	
[Expenses]	LOAD[hours]*[Hour]
[Load]	SUM(LOAD[hours])
[Day and month]	DAY(STATUSES[Date]) & "." & MONTH(STATUSES[Date])

In a multidimensional model, apart from measures and calculated columns, dimension hierarchies are being built. For instance, the department is responsible for implementing the training program, while the department itself is part of an institute structure, resulting in "Institute → Department → Training Program" hierarchy.

3 Research Results and Discussion

In Business Intelligence, descriptive analytics is the first step, which is followed by visualization of results. In this context, visualization is a set of methods of presenting initial information and data analysis results in a form that is easy to read and interpret, i.e. graphical. In addition, visualization can be used to monitor the construction process and performance of various analytical models, verifying hypotheses and for other purposes,

related to data analysis. Power BI Desktop in conjunction with Power BI Report Server enables one to create and place interactive dashboards – tools that visualize and analyze data, demonstrate how a company's key indicators change over a certain period. Thus, based on the created multidimensional data model, a series of dashboards were developed, with the most significant one presented below. In order not to disclose confidential data on the university and its staff, figures will be distorted.

3.1 "University Economic Performance" Dashboard

"University economic performance" dashboard is used to visualize and analyze data in terms of training program implementation costs (Fig. 5). On the left, the "Institute → Educational program → Form of study" is presented. Measures are created in such a way that while moving from one element of the hierarchy to another, drill-up and drill-down operations are performed to generalize or, conversely, refine data. The central part of the dashboard shows the following:

- total expenses for training program implementation, in millions of rubles (calculated based on the sum of expenses for the implementation of the curriculum and general expenses);
- university costs in terms of training one student under the program;
- number of professors implementing the program;
- total income from students on budgetary and contractual (fee-paying) basis;
- income from one student in the chosen training program;
- number of students;
- marginality (see Table 2);
- planned gross margin, in millions of rubles;
- expected gross margin, in millions of rubles;
- arrears (fee-paying students' overdue payments);
- actual gross margin, in millions of rubles.

If the value is negative in the Margin column (Fig. 5, training program index 45.03.01), the value is highlighted as red. It means that this training program is unprofitable for the university, and the expenses for this program exceed the income received from enrolling students on contractual and state-funded basis. In order to understand what has led to such a situation, it is necessary to examine the other columns in the table. The example shows that the number of part-time students is too small, although the university's staffing costs are the same and not contingent on the number of students. In this situation, the management may take various managerial decisions (increasing the number of students, closing the training program, reducing staffing costs by lowering their qualifications, etc.).

In the lower part of the dashboard there are interactive pie charts that allow filtering and comparing total incomes and expenses per institute. In addition, in the right part of the dashboard there is a filter that enables displaying data according to study years and forms of education.

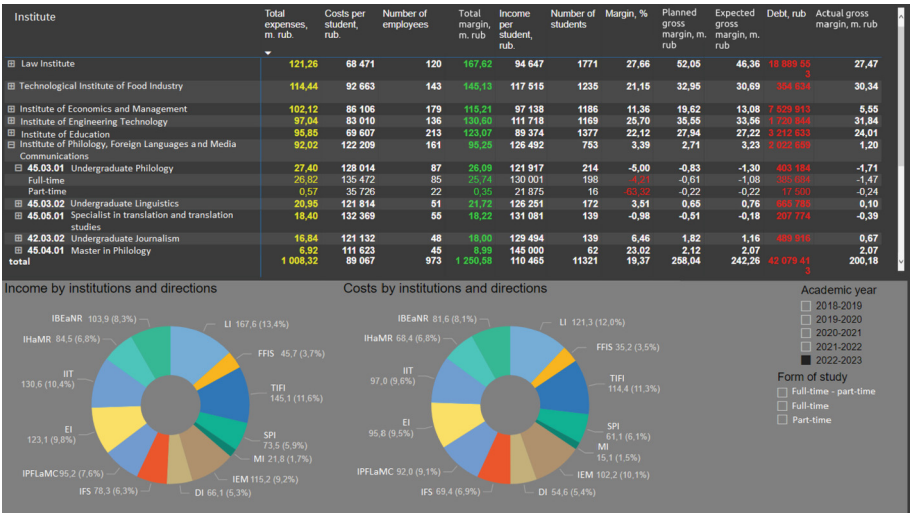


Fig. 5. "University economic performance" Dashboard.

3.2 Dashboards "Income from the Budget Form of Financing" and "Income from the Contract Form of Financing"

"Income from the budget form of financing" dashboard is used to visualize and analyze data on the university's budgeting from enrolling applicants on a state-funded basis (Fig. 6). On the left of the dashboard one can see a two-level hierarchy "Institute → Direction". By choosing a specific institute or training program at KemSU, all the other elements of the dashboard are automatically rebuilt and show either total data on the institute or specific data for the training program. The blue histograms in the upper part of the screen (Fig. 6) display the budget in millions of rubles and the target enrollment figures issued by the state to the university in different years. The budget is calculated based on the cost of education for one applicant in a particular training program established in a particular year (not all programs cost the same). In this case, the histograms contain 4 columns, because the dashboard provides data on the students taking a Bachelor's degree, which is 4 years long. The histograms allow one to evaluate volume dynamics for allocated funds and budget places for the university. If you simultaneously choose a specific institute or training program, it becomes possible to assess the dynamics by institute or program.

If the university did not expel students for poor academic performance, the planned income for September 1 of the current year would be equal to the sum of the values for the years from the blue histogram. However, due to the fact that deductions for academic failure can occur during any year of study, the closer a student is to graduation, the fewer classmates he or she has—hence the smaller the university's budget. In addition, the expulsion mechanism for academic failure does not imply instant expulsion in the event of an unsuccessful examination or test. The student is given time to retake the exam, allowing them to either reduce or accrue even more academic debt.

Fig. 6. Dashboard "Income from the budget form of financing".

The situations described above (the student's academic debt) must be modeled in order to predict budget losses in the event of a student being expelled. Thus, a decision was made to divide all students into 4 groups: students who successfully pass all the exams (such students do not lead to budget losses, therefore they are not displayed in the dashboard); students who have from 1 to 3 academic debts (the green column of the histogram at the bottom parts of the screen); students who have from 4 to 7 (yellow column) and more than 7 academic debts (red column). The last two categories of students are the most likely candidates to be expelled for academic failure, which means that it is of utmost importance to track such students amidst the entire student body at an institute and predict losses. The value on the right side of the "Expected income with dropouts" dashboard displays the predicted budget value in the case when students who fall into the second, third and fourth groups are expelled for poor academic performance (green, yellow and red columns). An interactive bar chart allows one to predict losses by clicking on a specific column. It has been established in practice that while students from the second group (green column) have high chances of continuing their education, the chances for the ones from the fourth (red) column are extremely low. In addition, by selecting a specific column of the histogram, you can determine the institutions and directions in which these students study. Directors of institutes, being able to see the current situation, can make various managerial decisions to keep it (schedule retakes, additional consultations, etc.). As one can see, compared to the previous year, the work carried out by the directors of the institutes resulted in the number of unsuccessful students with 4–7 and more than 7 debts decreasing and the number of unsuccessful students with 1–3 debts increasing, meaning that there are fewer students who failed their exams, which has a net positive effect on the university's budget.

The "Income from the contract form of financing" dashboard provides the same functionality (Fig. 7).

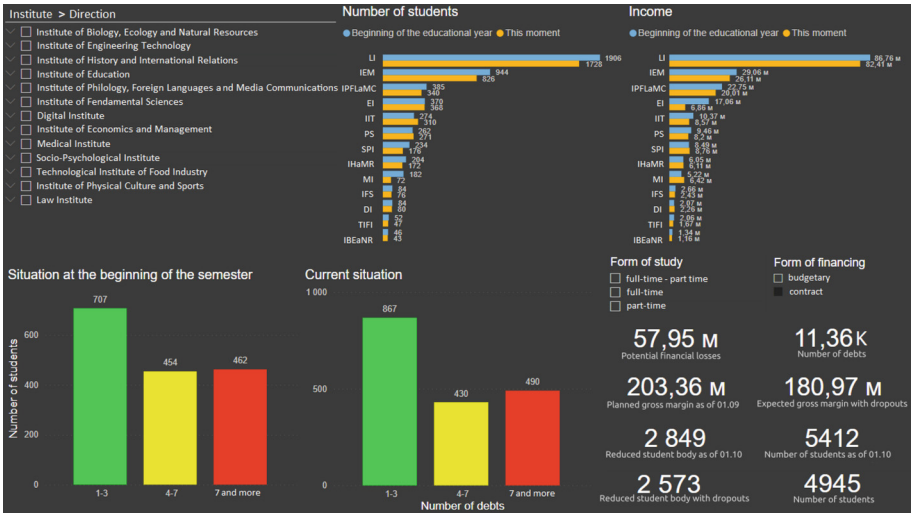


Fig. 7. Dashboard "Income from the contract form of financing".

The only difference here is that the number of students, enrolled on a contractual basis is impossible to predict – it can only be known a posteriori, i.e. when the students are actually enrolled.

4 Conclusion

This paper presents an approach to analyzing a university's economic performance using modern BI technologies. This approach allows reducing labor costs, quickly monitoring the current and projected values of the university budget, finding problem areas, testing emerging hypotheses and making managerial decisions that can lead to improved economic performance. Although from the point of view of economic science such an analysis is not comprehensive and justified, in addition, there are no standards defined by the Ministry of Science and Education (regulator) that can act as a Key Performance Indicator to assess the economic efficiency of the implementation of areas of study. It nonetheless enables one to obtain data quickly, which is advantageous, giving time and opportunity to rectify the situation and save student enrollment, which will ultimately lead to the preservation of budget funds by the university. In case of approaching the threshold value (such a threshold exists, introduced by the founder and expressed as a percentage of the reduced student) of student losses, the dashboard will display this dangerous situation in red!

Analysis of the economic feasibility of a particular training program is difficult without the use of such BI systems. The difficulty lies in correlating financial costs to a particular program, because, for instance, the same educator can work in different areas of training, and his or her remuneration is calculated only from the number of classes conducted without taking into account the program, necessary qualifications or material base. Implementing the created Decision Support System at the university allows one to

carry out a more thorough analysis of the implementation of the training programs that display negative values in the marginality column (Fig. 5). In such areas, decisions can be made to increase the enrollment of applicants, which can lead to an increase in the revenue side of the university's budget with a fixed expenditure.

The results of this study, with the necessary adaptation of ETL processes, can be applied to other universities in the country if the following conditions are met:

– The university's budget financing model is standard for most universities and is based on the enrollment target figures and the rules for the phased allocation of budgetary funds with control over the non-exceeding of the threshold value of student losses.
– The university has a sufficiently developed information infrastructure capable of providing the data sources necessary for analysis.

The research was conducted on the equipment of the Research Equipment Sharing Center of Kemerovo State University, agreement No. 075-15-2021-694 dated August 5, 2021, between the Ministry of Science and Higher Education of the Russian Federation (Minobrnauka) and Kemerovo State University (KemSU) (contract identifier RF----2296.61321X0032).

References

1. Moscoso-Zea, O., Castro, J., Paredes-Gualtor, J., Luján-Mora, S.: A hybrid infrastructure of enterprise architecture and business intelligence & analytics for knowledge management in education. IEEE Access **7**, 38778–38788 (2019). https://doi.org/10.1109/ACCESS.2019.2906343
2. Hamoud, A.K., Hussein, M.K., Alhilfi, Z., Sabr, R.H.: Implementing data-driven decision support system based on independent educational data mart. Int. J. Electr. Comput. Eng. **11**(6), 5301–5314 (2021). https://doi.org/10.11591/ijece.v11i6.pp5301-5314
3. Piedade, M.B., Santos, M.Y.: Business intelligence in higher education: enhancing the teaching-learning process with a SRM system. In: 5th Iberian Conference on Information Systems and Technologies, Santiago de Compostela, Spain, pp. 1–5. IEEE (2010)
4. Charikov, P.N., Fokin, S.: Study of the student progress monitoring system of a university department for a distance learning system based on a Web portal and for OLAP samples using to automate analysis. J. Phys. Conf. Ser. **1515**, 022073 (2020). https://doi.org/10.1088/1742-6596/1515/2/022073
5. Zina, A.S.A., Obaid, T.A.S.: Design and implementation of educational data warehouse using OLAP. Int. J. Comput. Sci. Netw. **5**(5), 824–827 (2016)
6. Savina, A.G., Malyavkina, L.I.: Business intelligence technology as a tool of big data analysis and support of administrative decision making application (In Russian). OrelSIET Bull. **1**(55), 85–92 (2021). https://doi.org/10.36683/2076-5347-2021-1-55-85-92
7. Karabtsev, S., Kotov, R., Davzit, I., Gurov, E.: Support decision-making in the management of the educational institution contingent based on Business Intelligence (in Russian). Prikladnaya informatika J. Appl. Inform. **17**(5), 125–142 (2022). https://doi.org/10.37791/2687-0649-2022-17-5-125-142
8. Larson, B.: Delivering Business Intelligence with Microsoft SQL Server 2016, 4th edn. McGraw-Hill Education, New York (2016)
9. Larson, B.: Data Analysis with Microsoft Power BI. McGraw-Hill Education, New York (2019)
10. Inmon, W.H.: Building the Data Warehouse, 4th edn. Wiley, Indianapolis (2005)

11. Suman, S., Khajuria, P., Urolagin, S.: Star schema-based data warehouse model for education system using Mondrian and Pentaho. In: Goel, N., Hasan, S., Kalaichelvi, V. (eds.) MoSICom. LNEE, vol. 659, pp. 30–39. Springer, Singapore (2020). https://doi.org/10.1007/978-981-15-4775-1_4
12. Kimball, R., Ross, M.: The data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling, 3rd edn. Wiley, Indianapolis (2013)

Explainability Improvement Through Commonsense Knowledge Reasoning

HyunJoo Kim and Inwhee Joe(✉)

Hanyang University, 222 Wangsimni-ro, Seongdong-gu, Seoul, Korea
{hjkim666,iwjoe}@hanyang.ac.kr

Abstract. The explainable artificial intelligence (XAI) studies on the black box model have forcused on a number of area such as feature importance, model agnostic studies, and surrogate model methods. XAI is necessary to increase the explanatory power of feature importance of data and contribute to improving the performance of models. Using human common sense in XAI makes it easier for humans to understand, but such research is lacking. In this paper, we propose a commonsense-learned model and reasoning process to obtain explanatory power which can explain parts that the model could not explain previously in structured data. We extracted common sense about age from ChatGPT, which has recently become a hot topic. Commonsense was used for preprocessing and interpretation of the model results to increase explanatory power and help to understand features. The explanatory power of the model was expressed by Shapley additive explanations and local interpretable model-agnostic explanations and this contributed to the fact that local data could be explained using a commonsense approach learned by humans. *abstract* environment.

Keywords: XAI · Commonsense reasoning · knowledge reasoning · Black box · Explainable AI

1 Introduction

Explainable artificial intelligence (XAI), which enables humans to understand the results and decisions of artificial intelligence (AI) such as the degree of trust and the reason for errors, is being studied in many fields. For example, local interpretable model-agnostic explanations (LIME) can be applied regardless of the artificial intelligence model by finding an explanation after the result [1]. This approach is a way to observe how the output changes by changing the input little by little. This reveals which inputs have the greatest impact on the output. In addition, in the case of the Shapley additive explanations (SHAP) which is a unified approach to interpreting model predictions [2], the SHAP value is used, this method calculates the contribution of each player in each game based on game theory. Therefore, SHAP stands for a conditional expectation of the Shapley value. Deep SHAP [3] is a method of calculating a description

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 259–277, 2024.
https://doi.org/10.1007/978-3-031-53552-9_24

of a model by combining SHAP and DeepLIFT [4]. Taking a simple neural network as an example, it receives $y = (y1, y2)$ as an input, passes through the hidden layers corresponding to $f1$ and $f2$, and finally calculates $f(y)$. The DeepLIFT method fixes the output from this simple model to 1 and calculates what contribution score the nodes of each layer have to produce an output corresponding to 1. In addition, the DeepSHAP method derives the degree of influence of the characteristics by obtaining contribution scores for f(y) of $x1$ and $x2$ calculated from $f1$ and $f2$ corresponding to the layers immediately before entering the output layer $f3$.

These technologies explain the results of artificial intelligence by using the results of models created with characteristics and outputs when calculating the importance and influence of characteristics. Currently, AI has developed in various fields, and the progress is remarkable. However, to ultimately become strong artificial intelligence, it will be necessary to think like humans to improve performance. To this end, research is also being conducted to derive common sense, a reference that humans already know by learning symbolically. Research on learning pre-trained models using the results after common sense reasoning [5] has also been conducted, and research on deriving common sense from pre-trained-models [6] has also progressed considerably. In this study, we propose a method to increase the explanatory power of XAI by utilizing common sense in addition to existing explanation methods. To increase the explanatory power of the feature importance and contribution of the model in structured data, the reasoning was performed using a machine learning model trained with commonsense data. Commonsense about age is extracted from ChatGPT-based GPT3 [7]. Through the reasoning process, we have the explanatory power to explain the part that the model could not explain in the past, but the performance was not improved by increasing the performance accuracy of the model. In SHAP and LIME, the explanation of the model technique was presented, and this helped explain local data using common sense modeled on human reasoning.

2 Related Work

2.1 Explainable AI

Explainable artificial intelligence refers to a technology that enables humans to understand, correctly understand and interpret the operation and final results of AI, as well as to explain the process by which the results are created. Currently, the most popular algorithm in AI is the neural network [8]. The neural network has a good performance, but the developer or user does not know how or why the AI produces such results because of a structure beyond human cognitive ability; therefore, AI is often called a block box problem. Despite its relatively short history, many techniques have been studied to explain AI, and there are three criteria for classifying XAI [9]. As described in Table 1, each criterion has a different perspective, but each viewpoint is not hierarchical, XAI classification does not follow a single viewpoint.

Table 1. XAI Category

View Point	Category
Complexity	Intrinsic vs Post-hoc
Scope	Global vs Local
Dependency	Model-specific vs Model-agnostic

2.2 Complexity

The complexity of a model is closely related to its explanatory power. The more complex the model, the more difficult it is for humans to understand or interpret. Conversely, the simpler the model, the easier it is for human comprehension. However, since a complex structure is advantageous for solving more difficult problems, there is a trade-off relationship between model complexity and interpretability. Neural networks gained great popularity because of their extremely high accuracy despite their low explanatory power, and their adoption is rapidly spreading to fields that do not require explanation.

Intrinsic. The fastest way to make a model easily interpretable is to devise an interpretable structure from the beginning. A simple model, such as a decision tree, is easy for humans to interpret just by looking at its structure. A model with a simple structure is called intrinsic because its already secured interpretability is self-evident, or it is also expressed as having transparency. The strength of how intrinsic it is can be answered by the question, "How does the model work?". However, because of this trade-off relationship, intrinsic models have low accuracy. Linear models can also be classified as intrinsic. In a linear model [10], the coefficient of the independent variable indicates the importance of the independent variable. Alternatively, if L1 regularization is applied, it becomes a sparse linear model in which the coefficients of some independent variables are 0.

Post-hoc. If the model itself does not have explanatory power, the predicted results of the model can only be interpreted post-hoc. Most interpretable techniques in the field of machine learning and deep learning are post-hoc. It would be best if the model had a high accuracy and inherent explanatory power, but this is difficult to exist and achieve in practice. It is common to use a complex model with good performance but post-hoc analysis.

2.3 Scope

Interpretation techniques can be classified according to the scope described. They are divided into a global technique that always has explanatory power for all prediction results and a local technique that can explain only one or some prediction results.

Global. The global method describes all the outcomes predicted by the model based on the understanding of the model's logic. Or, at the module level, it also includes the scope to describe how one module of the model affects the predicted outcome. The intrinsic model belongs to the global technique by nature because it can explain all the prediction results from the structure of the model. Models such as decision Trees, and falling Rule lists [11] belong to this category. Global is an ideal technique, but it is really difficult to implement post-hoc as global. In terms of explanation, even if a certain explanatory power is available for all predictions, the ability to explain the characteristics of individual prediction results may be somewhat poor.

Local. Local techniques account for only a specific decision or one predicted outcome. The scope of grouping several prediction results and explaining the prediction group is also included in the local technique. Compared to the global method, the local method has less scope to explain, so it is relatively feasible and costs less. Although it cannot explain the overall prediction propensity, one or a few prediction results can be explained almost perfectly. Realistically, every prediction does not require an explanation. It is realistic to explain the predictions where the issue arose. It is more practical than the global method in that it can at least explain the issue well when an explanation is needed. Local Interpretable Model-agnostic Explanation (LIME), a representative XAI methodology, can be classified into Local from the Scope point of view, model-agnostic from the dependency point of view, and post-hoc from the Complexity point of view. Important input variables can be found by perturbing the inputs to see how the predictions change.

2.4 Dependency

Dependency can be classified according to whether the explanation technique is specialized to apply only to a specific model or whether it can be universally applied regardless of the model [12].

Model-Specific. Models can be applied, which reduces the choice when choosing a model. For example, all visualization analysis techniques that can only be used in CNNs are model-specific. For example, the activation method explains the model by visualizing the feature map or weight itself. It is model-specific and global at the same time. It is also a post-hoc technique because it is explained after the model is created. In addition, when finding an element that has contributed to CNN, the saliency map expresses the part that can intuitively understand where the model looked was predicted by displaying the location of the element on the image [13]. In addition, the class activation map (CAM) [14] visualizes which part of the image was seen and predicted in the image classification model. Instead of flattening from fully connected CNN, we use Global Average Pooling to identify which part contributed to the class classification.

In the case of the gradient class activation map(Glad-CAM) [15], which complements the disadvantages of recent CAM, the model structure is left as it is, and the gradient for the feature map channel is obtained for the weight for each channel.

Model-Agnostic. In cases that are model-agnostic, in which the inside of a model is unknown to humans, we must find evidence outside the model for an explanation as it does not use any characteristics of the model. Therefore, some features can be applied regardless of the model. Representative methodologies include Surrogate [12], sensitivity analysis [16], partial dependence [17], partial dependency plot (PDP) visualization, and individual conditional expectation (ICE) [18] methods. The example based on [19] as an explanation method, used a technique to explain the model by selecting a specific example from the dataset. The representative one is the maximum mean discrepancy-critic (MMD-critic), which is a technique to automatically find prototypes and criticism in an unsupervised learning method [20].

2.5 Commonsense Reasoning

NYU's Professor Ernest Davis characterized common sense knowledge as "what a typical 7-year-old knows about the world." This includes physical objects, substances, plants, animals, and human societies. In general, book learning, expertise, and knowledge of customs are excluded. However, sometimes knowledge of these subjects is included. For example, knowing how to play cards is expertise, not common sense. However, it is considered common sense to know that people play cards for fun [21,22]. In AI, commonsense reasoning is the human-like ability to make inferences about the types and nature of everyday situations that humans encounter daily [23]. These assumptions include judgments about the characteristics of physical entities, their taxonomic properties, and their intentions. Similar conclusions can be drawn from our innate ability to reason about human behavior and our natural understanding of the physical world. Methods of inference of commonsense knowledge from NLP pre-trained models [24–26] have been attempted. ChatGPT is a kind of natural language processing model that has been developed using deep learning techniques. The architecture of ChatGPT is based on the transformer model GPT3 [7], which is a type of neural network for natural language processing tasks. ChatGPT has been pre-trained on a massive amount of text data, which allows it to generate language responses to a wide range of questions and prompts. There are suspicions of pointers, deterred by commonsense from the current pre-trained-language model(PLM). Better explanatory results can be obtained if PLM receives common sense from socially and culturally recognized product data. Common sense reasoning in PLM (chatGPT) is a catch-up method that can overcome the explanatory power of XAI in terms of methodology There are various studies related to Capturing Commonsense Knowledge and Reasoning with Commonsense Knowledge. Linguistic-related research [27], semantic similarity-related research [28], research using

Attention maps [29], numerical reasoning-related research [30], etc. There are various studies that try to reason commonsense from PLM, and ChatGPT is one of these PLM.

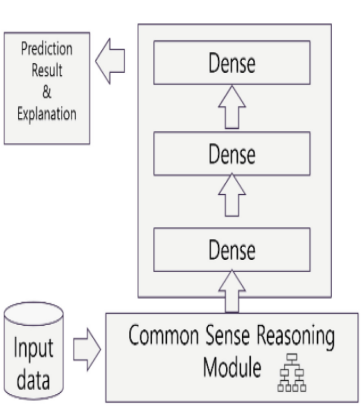
3 Methodology

3.1 Problem Statement

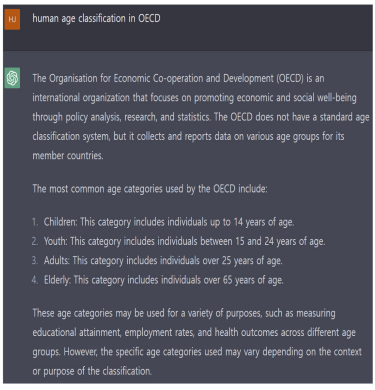
Neural networks perform well, but their complex mathematical structure makes it difficult to interpret AI results. This dilemma is a typical black box problem. Therefore, most studies on neural networks in the XAI area have been confined to visual representations of image data based on CNNs. NLP-related XAI research is also being conducted to show the basis for answering questions in data such as XQA, but there have not been many XAI research attempts in standardized data. This paper focuses on XAI research through commonsense reasoning in structured data with categorical and numeric features.

3.2 Proposed Methodology

This methodology first learned a common sense model trained with data having commonsense information. It derived explanatory power between the features entering the inside of the neural network through reasoning with the commonsense model in the neural network. The learning method and architecture were as follows:



(a) Model Architecture



(b) Commonsense for age from ChatGPT.

Fig. 1. Commonsense-Learning dataset example and model architecture.

The main model was designed as a simple neural network. The structure consisted of a Neural Network with three layers (Dense-Dense-Dense/Sigmoid), and

Adam was used as the optimizer. Figure 1a shows the architecture of the model. The Common Sense Reasoning module is composed of a decision tree trained with common sense for age as a model. The input data was passed through the commonsense reasoning module, and the result after reasoning was trained by the neural network. Our neural network was trained to solve a classification problem. A loss function is a function that compares the target value and predicted values. It is trained to minimize this loss function between predicted and target outputs. This formula Eq. 1 is as below:

$$J(W^T, b) = \frac{1}{m} \sum_{i=1}^{m} L(\hat{y}^{(i)}, y^{(i)}) \tag{1}$$

where w^T and b are the weights and biases that minimize the value of J (average loss)

$$L = \frac{1}{N} \left[\sum_{j=1}^{N} [t_j \log(p_i) + (1 - t_j) log(1 - p_j)] \right] \tag{2}$$

Equation 2 is for the binary classification loss function. So cross-entropy is calculated as average cross-entropy for all data examples. Figure 1b is commonsense data. It is extracted from ChatGPT trained language model with massive data. Generally, we are supposed to be commonsense for age from this result. The core algorithm for commonsense reasoning is decision trees by ID3 [31] which employs a top-down, greedy search using entropy. This method uses Standard Deviation Reduction to make a decision tree for regression.

$$n = count\ of\ data \tag{3}$$

$$Average = \bar{x} = \frac{\sum x}{n} \tag{4}$$

$$Standard\ Deviation = s = \sqrt{\frac{\sum (x - \bar{x})^2}{n}} \tag{5}$$

$$Coefficient\ of\ variation = \frac{s}{\bar{x}} * 100\% \tag{6}$$

N is for a count of data like Eq. 3. Equation 4 is the mean of x. Training is for reducing the standard deviation Eq. 5 after a dataset is split on an attribute. The coefficient of variation is calculated as Eq. 6. Our attribute is one column for the reasoning result of age as in Fig. 1b. If a dataset such as commonsense that can be organized in this way exists, various reasoning is possible and the performance and explanatory power of the model can be increased.

3.3 Dataset Description

The following datasets have been used in our experimentation, the Titanic dataset, Bank Marketing dataset [32], in-vehicle coupon recommendation dataset, and Adult dataset.

Titanic Dataset. The Titanic dataset describes the survival status of individual passengers on the Titanic.

Bank Marketing Dataset. In the Bank Marketing Dataset, the data is related to direct marketing campaigns of a Portuguese banking institution. The marketing campaigns were based on phone calls. Often, more than one contact with the same client was required, in order to access if the product (bank term deposit) would be ('yes') or not ('no') subscribed. The data includes information on 45211 clients. Each record has 17 attributes. This dataset contains 10 columns of categorical data and 7 columns of numerical data.

In-Vehicle Coupon Recommendation Dataset. The In-Vehicle coupon recommendation dataset [33] was collected via a survey on Amazon Mechanical Turk. The survey describes different driving scenarios including the destination, current time, weather, passenger, etc., and then asks the person whether he will accept the coupon if he is the driver. Each record has 26 attributes composed of 16 columns of categorical data and 10 columns of numerical data.

Adult Dataset. The extraction of the Adult dataset was done by Barry Becker from the 1994 Census database. A set of reasonably clean records was extracted using the following conditions: $(AAGE > 16) \cap (AGI > 100) \cap (AFNLWGT > 1) \cap (HRSWK > 0)$. The aim is to determine whether a person's salary exceeds \$50,000 USD. The data includes information on 48842 person. Each record has 14 attributes composed of 8 columns of categorical data and 6 columns of numerical data.
Data preprocessing is like this:
- Numerical Features: It was processed and used using standard normalization.
- Categorical Features (Encoding): Label Encoding of the scikit-learn package was used. - Labeling: The label was encoded as 0/1.

4 Result and Discussions

4.1 Experimental Design

We used the Titanic dataset that traditionally solves binary classification problems. A binary Classifier was used as the main model, and a decision tree was used as the model for reasoning. Let us call a model with common sense data a reasoning model. To measure performance, the accuracy, precision, and recall of the classifier model were measured, and the loss and accuracy graphs were provided in the learning stage.

4.2 Result

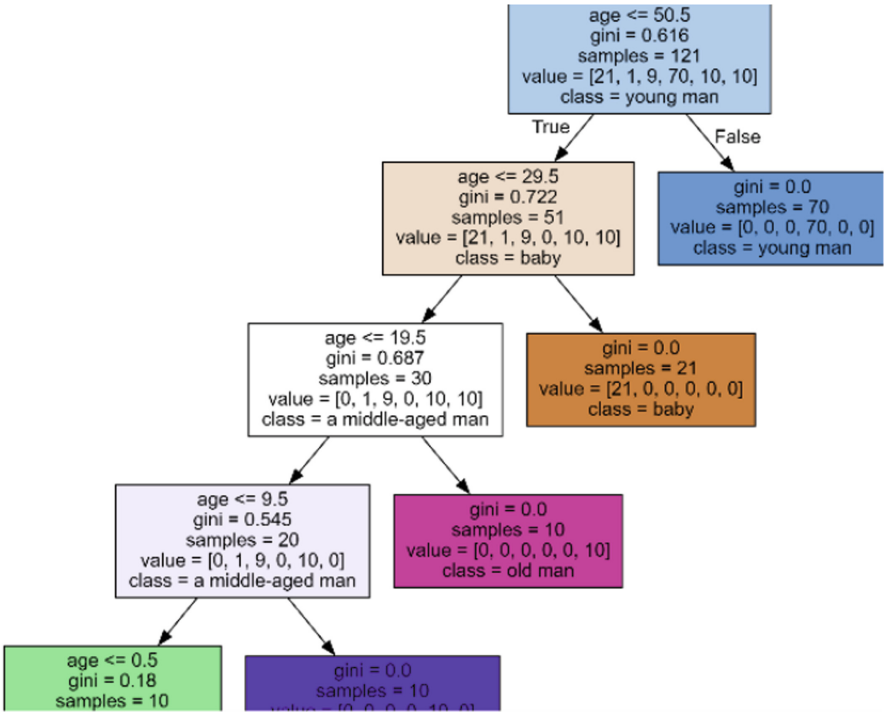


Fig. 2. Tree-based reasoning model (trained using commonsense data).

Figure 2 is a model that learned the common sense for age. It was trained with the decision tree model and shows the visualized results.

Figure 3 shows the visualization results of the LIME explanation for class survived. Red indicates a negative effect on the model, and a green bar indicates a positive effect. The features related to sex show the greatest influence, and the age factor of the kids, which is the result of reasoning, also has a great influence. As features that show a positive effect, it appears that the features of fare and class_first have an effect on survival.

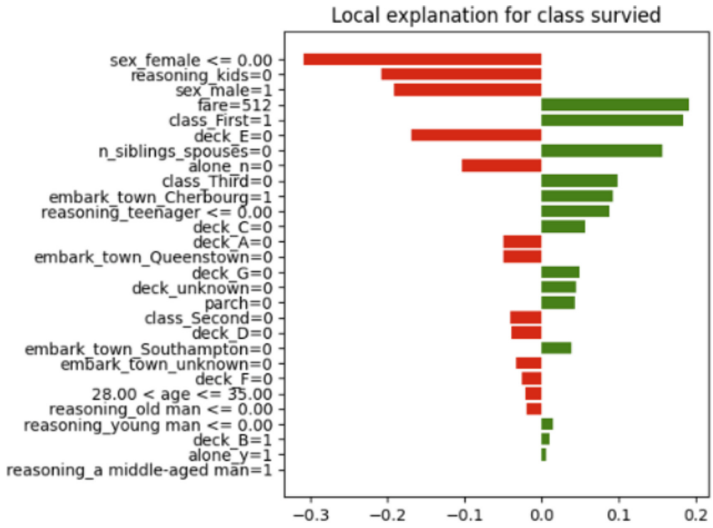


Fig. 3. Titanic dataset explanation using LIME after reasoning

Figure 4 shows no reasoning. Compared to Fig. 3, In the case before reasoning, the number of explainable features is greater than no reasoning and the variance of impact is smaller than in the case after reasoning. The larger the variance of the impact, the more clearly it can be explained between features.

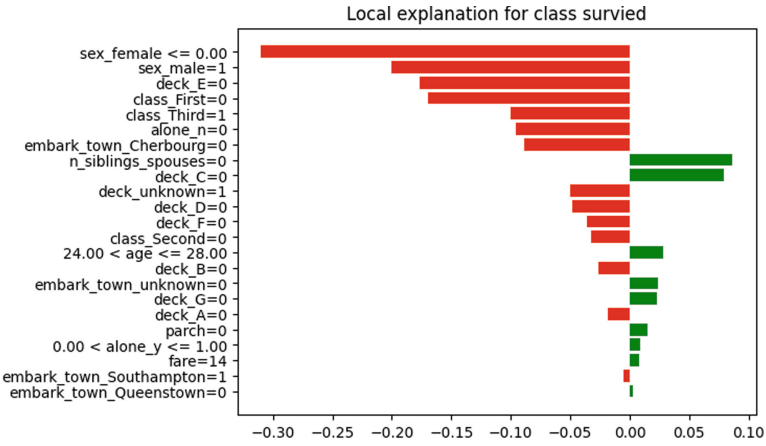


Fig. 4. Titanic dataset explanation using LIME without reasoning

Figure 5 shows the vehicle coupon recommendation dataset. The probability that the value of Y is 0 is 0.52, and the probability that it is 1 is 0.48. Here, the feature that influenced the prediction of 0 is *expiration* <= 0.00, *weather* <=

1.00, *age* <= 21.00, *income* and *marialStatus*. The features that influenced the prediction of 1 are *coupon* > 4, *destination* < 1, *toCoupon_GEG*15*min*.

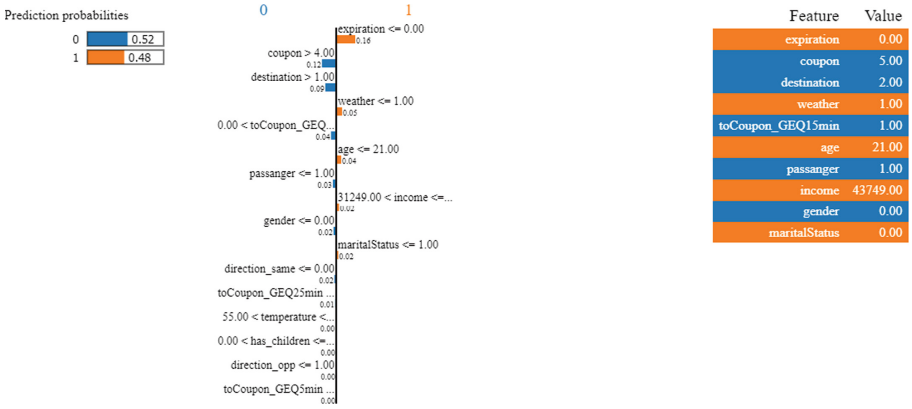


Fig. 5. Vehicle coupon recommendation dataset explanation using LIME without reasoning

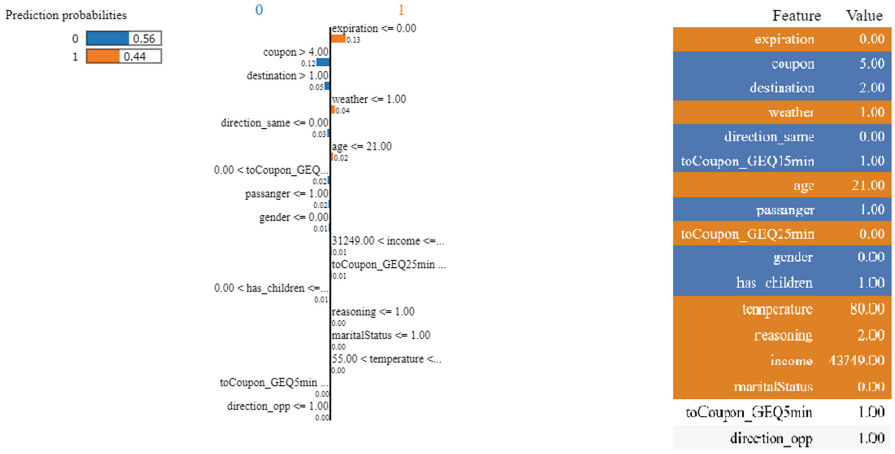


Fig. 6. Vehicle coupon recommendation dataset explanation using LIME after reasoning

Figure 6 shows the vehicle coupon recommendation dataset. It is the result of applying Reasoning to, and the probability of predicting Y as 0 is 0.56. The probability of predicting 1 is 0.44. The features that influenced the prediction as 0 are *coupon* > 4, *destination* > 1, *direction* <= 0, *toCoupon_GEG*15*min*, *passenger*, *gender* and *haschildren*. The feature that influenced the prediction of 1 is *expiration* < 0.0, *weather* <= 1.0, *age* <= 21.0, *toCoupon_GEG*15*min*,

temperature, *reasoning*, *income*, *marial_status* has influenced. After reasoning, the reasoning result is a feature. It acts as a factor affecting Y.

Table 2 shows that it is more clearly explainable between features after reasoning than before reasoning. The number of explanatory features before and after Reasoning (Yes/No) for the four datasets, Absolute value average for feature influence derived by Lime, derived by Lime The variance of the influence of the feature and the influence derived by Lime Standard deviations were tabulated. Lightgbm was used as the model. Before and after reasoning were compared. In the case of Titanic data, the number of features was 9 before reasoning. After reasoning, the number of features increased to 10, which is the number of explainable features. It shows an increase the absolute value average. It increased from 0.0212 to 0.0409 after reasoning. In terms of impact, it shows information that has a greater impact. The variances went from 0.0027 before reasoning to 0.0087 after reasoning increased, which is more diverse and has a large or large variance in the influence, so the explanatory power is it means more clear. Std also at 0.00518 before reasoning Later, the value increased to 0.0935. Similarly, of the described feature The larger the variance, the greater the explanatory power.

In the case of the vehicle coupon dataset, the number of explainable features is reasoning It was 16 before, and increased to 17 after reasoning. The number has increased. The absolute mean was 0.038 before reasoning and 0.0281 after. It did not show a significant difference, and variance and Std also did not show a significant increase. For the adult income dataset, the number of explainable features is It was 10, but later, the explanatory features increased to 11. The absolute mean increased from 0.0897 before reasoning to 0.107 after, and Variance also From 0.0405 to 0.0239, it can be seen that the variance in feature explanatory power has increased. It can be seen that Std also has a clearer explanatory power as it increases from 0.1544 to 0.2012. In the case of the bank marketing dataset, explainable features increased from 8 to 9 after reasoning, and the absolute mean increased from 0.062 to 0.1154. Variance also increased from 0.0097 to 0.0362, and Std also increased from 0.0986. It can be seen that the explanatory power has increased by increasing the deviation to 0.1902. It can be seen that the explanatory power increased in three of the four datasets.

Table 3 uses lightgbm as a classification model. The performance of the model before and after reasoning was compared. In the case of the Titanic dataset, the accuracy was 0.8317 before reasoning. After reasoning, it showed a slight decrease to 0.7619, and the precision increased from 0.8393 to 0.9412. Recall decreased from 0.7121 to 0.4267, and F1 also decreased from 0.7705 to 0.5872. In the case of the vehicle coupon dataset, the accuracy showed a slight performance decrease from 0.6894 to 0.6729, and the precision also decreased from 0.7037 to 0.6729. The recall increased from 0.7807 to 0.8053, F1 score also slightly decreased from 0.7402 to 0.7362. It is a very slight difference. The adult income dataset also shows very small differences. Accuracy decreased from 0.8869 to 0.8663, precision was from 0.7793 to 0.7756, recall also decreased very slightly from 0.6201 to 0.6218, The F1 score also decreased very slightly from 0.6902 to

Table 2. Comparison of LIME result using Reasoning or not with lightgbm

Dataset	Reasoning	Features	Mean	Variance	Std
Titanic	Yes	10	0.0409	0.0087	0.0935
	No	9	0.0212	0.0027	0.0518
Vehicle coupon	Yes	17	0.0281	0.0022	0.047
	No	16	0.038	0.0033	0.0571
Adult income	Yes	11	0.107	0.0405	0.2012
	No	10	0.0897	0.0239	0.1544
Bank marketing	Yes	9	0.1154	0.0362	0.1902
	No	8	0.062	0.0097	0.0986

0.6960. In the Bank marketing dataset, the accuracy is 0.7993 before reasoning. After reasoning, it increased to 0.8107, and the precision also increased slightly from 0.7913 to 0.7932. The recall also increased slightly from 0.7903 to 0.8189, and the F1 score also increased from 0.7908 to 0.8189. It can not be seen that reasoning greatly help to improve the performance of the model.

Table 3. Comparison of LIME result using Reasoning or not with lightgbm

Dataset	Reasoning	Features	Accuracy	Precision	Recall	F1
Titanic	Yes	10	0.7619	0.9412	0.4267	0.5872
	No	9	0.8217	0.8393	0.7121	0.7705
Vehicle coupon	Yes	17	0.6729	0.678	0.8053	0.7362
	No	16	0.6894	0.7037	0.7807	0.7402
Adult income	Yes	11	0.8663	0.7756	0.6218	0.6960
	No	10	0.8669	0.7793	0.6201	0.6902
Bank marketing	Yes	9	0.8107	0.7932	0.8189	0.8059
	No	8	0.7993	0.7913	0.7903	0.7908

Table 4 shows the number of explanatory features before and after reasoning (Yes/No) for the four datasets, absolute value average for feature influence derived by Lime, derived by Lime, the variance of the influence of the feature and the influence derived by Lime, standard deviations were tabulated. The model is a deep neural network It was used and compared before and after reasoning. In the case of Titanic data, the number of features was 9 before reasoning. After reasoning, the number of features increased to 10, which is the number of explainable features. The absolute value average increased from 0.0479 to 0.0766 after reasoning. In terms of impact, it shows information that has a greater impact. The variances increased from 0.0085 before reasoning to 0.0158 after reasoning. The large variance in the influence means the explanatory power improvement and

more clear. Std was also at 0.0923 before reasoning. Later, the value increased to 0.1257. Similarly, of the described feature The larger the variance, the greater the explanatory power.

In the case of the vehicle coupon dataset, it couldn't get result because model performance problem. For the adult income dataset, the number of explainable features was 10, but it increased to 11. The absolute mean decreased from 0.1675 to 0.1161 before reasoning, Variance also decreased from 0.1353 to 0.0811. Std also decreased from 0.3679 to 0.2849.

In the case of the bank marketing dataset, before reasoning, explainable features increased from 8 to 9 later, and the absolute mean ranged from 0.1667 to 0.1729. Variance also increased from 0.0573 to 0.0631, and Std also increased from 0.2394 to 0.2513. It shows that the explanatory power increased as the deviation increased. In two of the four datasets, it can be seen that the explanatory power increased after reasoning.

Table 4. Comparison of LIME result using Reasoning or not with deep neural network

dataset	Reasoning	Features	Mean	Variance	Std
Titanic	Yes	10	0.0766	0.0158	0.1257
	No	9	0.0479	0.0085	0.0923
Vehicle coupon	Yes	17	0.0000	0.0000	0.0000
	No	16	0.0000	0.0000	0.0000
Adult income	Yes	11	0.1161	0.0811	0.2848
	No	10	0.1675	0.1353	0.3679
Bank marketing	Yes	9	0.1729	0.0631	0.2513
	No	8	0.1667	0.0573	0.2394

Table 5 shows the result of a deep neural network as a classification model. The performance of the model before and after reasoning was compared. In the case of the Titanic dataset, the accuracy was 0.8217 before reasoning, it slight decreased to 0.71, and the precision decreased from 0.8393 to 0.6475, Recall is reduced from 0.7705 to 0.5844, and F1 is also decreased from 0.7705 to 0.5844. And AUC also decreased from 0.8182 to 0.7547. In the case of vehicle coupon dataset, due to poor model performance, we could not get the evaluation results before and after reasoning.

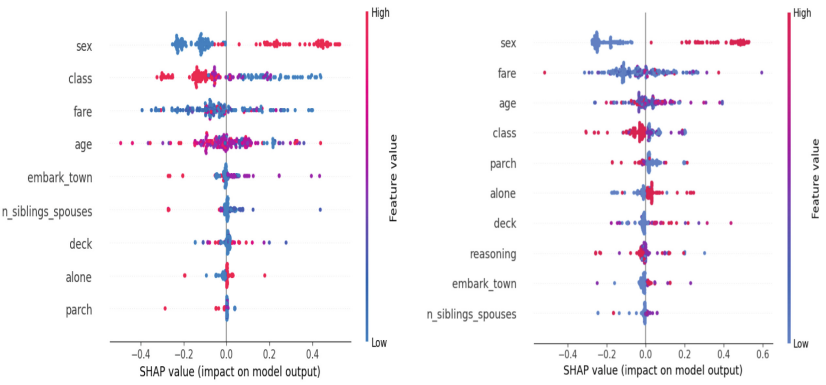
The adult income dataset also shows very small differences. Accuracy slight increase from 0.8017 to 0.8031, precision increased from 0.88 to 0.8091, recall also increased slightly from 0.2967 to 0.2412, F1 score also increased slightly from 0.3075 to 0.3443, AUC value decreased from 0.6316 to 0.6304.

Table 5. Comparison of LIME result using Reasoning or not with deep neural network

dataset	Reasoning	Accuarcy	Precision	Recall	F1	AUC
Titanic	Yes	0.71	0.6475	0.5357	0.5844	0.7547
	No	0.8217	0.8393	0.7121	0.7705	0.8182
Vehicle coupon	Yes	0.5691	0.5692	1	0.7214	0.5
	No	0.5692	0.5692	1	0.7214	0.5
Adult income	Yes	0.8031	0.8091	0.2412	0.3443	0.6304
	No	0.8017	0.88	0.2067	0.3075	0.6316
Bank marketing	Yes	0.7732	0.7172	0.8563	0.7758	0.8515
	No	0.7839	0.7802	0.7539	0.7586	0.8518

In the Bank marketing dataset, the accuracy is at 0.7839 before reasoning. After that, it decreased slightly to 0.7732, the precision also decreased slightly from 0.7802 to 0.7172, recall increased slightly from 0.7539 to 0.8563, F1 score also increased from 0.586 increased to 0.7786, AUC value also slightly decreased from 0.8518 to 0.8515, but it is a very small difference. It can be seen that reasoning does not greatly help improve the performance of the model.

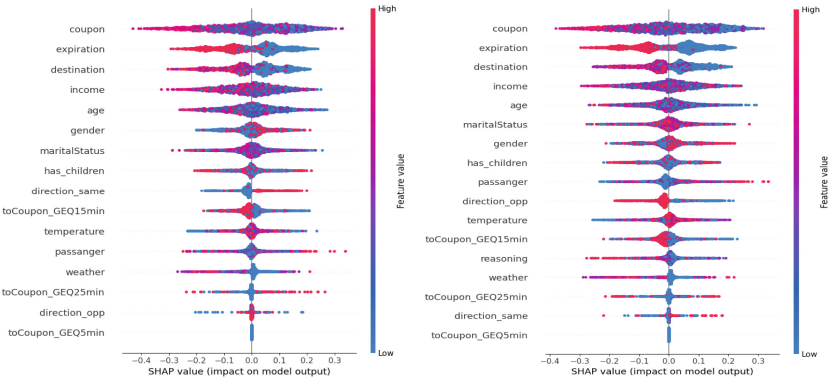
Figure 7 shows the SHAP values before and after reasoning on the Titanic dataset. Figure 7a is the extracted SHAP values without reasoning. The smaller the value of *sex*, the SHAP value The smaller the value and the larger the value, the larger the SHAP value. Male is labeled as 0 and female as 1. The gap is big In addition, *class*, *fare*, and *age* have large values and SHAP values. Influence of embark_town, *n_siblings_spouses*, *deck*, *alone*, *parch*, etc. on the model SHAP value is shown. In Fig. 7b, the SHAP result after reasoning is added with reasoning that was not in the existing feature. As the value of reasoning increases, the value of SHAP decreases, and as the value decreases, the value of SHAP increases.



(a) SHAP values without reasoning (b) SHAP values after reasoning

Fig. 7. SHAP values for titanic dataset

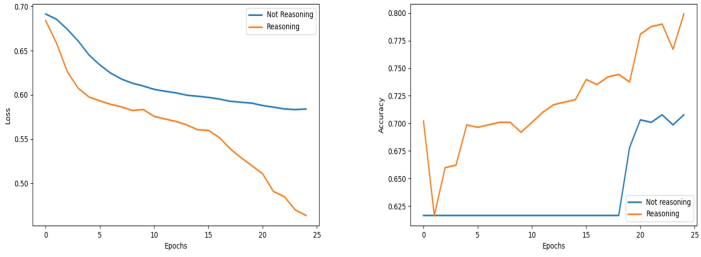
Figure 8 shows SHAP values before and after reasoning in the vehicle coupon recommendation dataset. Figure 8a is the extracted SHAP values without reasoning. In case of *coupon* feature, the larger the deviation of the feature value and the smaller the coupon value, the higher the SHAP value. SHAP value tends to decrease. Other than that, *expiration*, *destination*, *income*, *age*, *Passenger*, etc. show a large deviation of SHAP value depending on the value. In Fig. 8b, the SHAP result after reasoning is added with reasoning that was not in the existing feature. SHAP value is expressed. As the value increases, the value of SHAP decreases, and as the value decreases, the value of SHAP increases.



(a) SHAP values without reasoning (b) SHAP values after reasoning

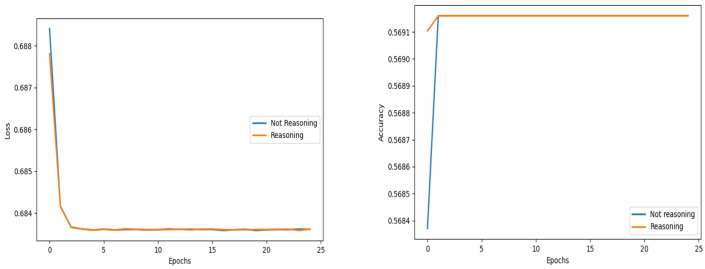
Fig. 8. SHAP values for Vehicle coupon recommendation dataset

Figure 9 is a graph plotting the changes for loss and accuary during the training process on the Titanic dataset. Changes when without reasoning and results after applying reasoning are shown in the same graph. Figure 9a is the training loss graph. Loss can be seen to drop faster after reasoning is applied. Figure 9b is the training accuracy graph. Accuracy increases faster after reasoning. Without reasoning the accuracy increased after epoch 25, but the result after reasoning showed better accuracy.



(a) Training loss (b) Training accuracy

Fig. 9. Model performance using reasoning or not for Titanic dataset.



(a) Training loss (b) Training accuracy

Fig. 10. Model performance using reasoning or not for Vehicle coupon recommendation dataset.

Figure 10 is a graph plotting the changes in loss and accuary during the training process on the Vehicle coupon recommendation dataset. Changes when without reasoning and results after applying reasoning are shown in the same graph. Figure 10a is the training loss graph. There is no significant difference between loss before and after reasoning. Figure 10b is the training accuracy graph. Accuracy increases faster after reasoning. However, as a result, there is no significant difference in accuracy.

The method of increasing the explanatory power of a neural network or lightgbm machine learning model after reasoning as a result of commonsense collected data is different from groundless feature engineering methods. It is different from the method of converting continuous data into categorical data through binning of the features. This is also different from the reason result through the data-based reasoning module as researchers arbitrarily divide the interval data.

5 Conclusions

In this paper, a method to increase the explanatory power of XAI was attempted by utilizing common sense in addition to the existing explanation method. To increase the explanatory power of the feature importance and contribution of the model in structured data, the reasoning was performed using a machine learning model trained with common sense data. Through experiments with four datasets, judging by the change in the influence value when Lime was applied to the model, 3 out of 4 datasets were derived that "reasoning" was meaningful. When comparing absolute mean, variance, and standard deviation values, except for the Vehicle coupon recommendation dataset, The numbers increased in the Titanic dataset, Adult income dataset, and Bank marketing dataset, respectively. In the performance of the model, as a result of comparing the evaluation results, no significant difference was found in the four datasets. Also, in the graph of model training, there was little difference in training loss and accuracy before and after reasoning. Reasoning did not affect the performance of the model. There were

results of reasoning in LIME feature effect value, SHAP value, and permutation importance, respectively, which can be seen as acting as a factor that increases the explanatory power of the model. Through the reasoning process, we have the explanatory power to explain the part that the model could not explain in the past, and the performance was not improved by increasing the performance accuracy of the model. The explanatory method of the model was expressed in SHAP and LIME and contributed to explaining local data using the common sense learned by humans.

Acknowledgement. This work was supported by the Institute for Information and Communication Technology Promotion (IITP) and funded by the Korean Government [Ministry of Science, ICT and Future Planning (MSIP)] (Development of the technology to automate the recommendations for big data analytic models that define data characteristics and problems) under Grant 2020-0-00107.

References

1. Tulio Ribeiro, S.S.M., Guestrin, C.: Local Interpretable Model-Agnostic Explanations (LIME): An Introduction. Oreilly, Newton (2016)
2. L.S.: Shapley. 17. A Value for n-Person Games, pp. 307–318. Princeton University Press, Princeton (1953)
3. Lundberg, S.M., Lee, S.-I.: A unified approach to interpreting model predictions. CoRR, abs/1705.07874 (2017)
4. Shrikumar, A., Greenside, P., Kundaje, A.: Learning important features through propagating activation differences. CoRR, abs/1704.02685 (2017)
5. Li, S., Chen, J., Yu, D.: Teaching pretrained models with commonsense reasoning: a preliminary KB-based approach. CoRR, abs/1909.09743 (2019)
6. Bhargava, P., Ng, V.: Commonsense knowledge reasoning and generation with pretrained language models: a survey. In: AAAI Conference on Artificial Intelligence (2022)
7. Brown, T.B., et al.: Language models are few-shot learners. CoRR, abs/2005.14165 (2020)
8. Hopfield, J.J.: Neural networks and physical systems with emergent collective computational abilities. PNAS (1982)
9. Ribeiro, M.T., Singh, S., Guestrin, C.: "Why should I trust you?": explaining the predictions of any classifier. CoRR, abs/1602.04938 (2016)
10. Clement, T., Kemmerzell, N., Abdelaal, M., Amberg, M.: XAIR: a systematic metareview of explainable AI (XAI) aligned to the software development process. Mach. Learn. Knowl. Extract. **5**(1), 78–108 (2023)
11. Wang, F., Rudin, C.: Falling rule lists. CoRR, abs/1411.5899 (2014)
12. Carvalho, D.V., Pereira, E.M., Cardoso, J.S.: Machine learning interpretability: a survey on methods and metrics. Electronics **8**(8), 832 (2019)
13. Xu, K., et al.: Show, attend and tell: Neural image caption generation with visual attention. CoRR, abs/1502.03044 (2015)
14. Zhou, B., Khosla, A., Lapedriza, À., Oliva, A., Torralba, A.: Learning deep features for discriminative localization. CoRR, abs/1512.04150 (2015)
15. Selvaraju, R.R., Das, A., Vedantam, R., Cogswell, M., Parikh, D., Batra, D.: Grad-CAM: Why did you say that? Visual explanations from deep networks via gradient-based localization. CoRR, abs/1610.02391 (2016)

16. Samek, W., Wiegand, T., Müller, K.-R.: Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models. arXiv preprint arXiv:1708.08296 (2017)
17. Krause, J., Perer, A., Ng, K.: Interacting with predictions: visual inspection of black-box machine learning models. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, pp. 5686–5697 (2016)
18. Goldstein, A., Kapelner, A., Bleich, J., Pitkin, E.: Peeking inside the black box: visualizing statistical learning with plots of individual conditional expectation. J. Comput. Graph. Stat. **24**(1):44–65 (2015)
19. Kenny, E.M., Ford, C., Quinn, M., Keane, M.T.: Explaining black-box classifiers using post-hoc explanations-by-example: the effect of explanations and error-rates in XAI user studies. Artif. Intell. **294**, 103459 (2021)
20. Kim, B., Khanna, R., Koyejo, O.O.: Examples are not enough, learn to criticize! Criticism for interpretability. In: Advances in Neural Information Processing Systems, vol. 29 (2016)
21. Davis, E.: Logical formalizations of commonsense reasoning: a survey. J. Artif. Intell. Res. **59**, 651–723 (2017)
22. Commonsense reasoning, December 2022
23. Davis, E., Marcus, G.: Commonsense reasoning and commonsense knowledge in artificial intelligence. Commun. ACM **58**(9), 92–103 (2015)
24. Lin, B.Y., Lee, S., Khanna, R., Ren, X.: Birds have four legs?! Numersense: probing numerical commonsense knowledge of pre-trained language models. arXiv preprint arXiv:2005.00683 (2020)
25. Feldman, J., Davison, J., Rush, A.M.: Commonsense knowledge mining from pretrained models. arXiv preprint arXiv:1909.00505 (2019)
26. Zhou, X., Zhang, Y., Cui, L., Huang, D.: Evaluating commonsense in pre-trained language models. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, pp. 9733–9740 (2020)
27. Bhagavatula, C., et al.: Abductive commonsense reasoning. CoRR, abs/1908.05739 (2019)
28. Niu, Y., Huang, F., Liang, J., Chen, W., Zhu, X., Huang, M.: A semantic-based method for unsupervised commonsense question answering. CoRR, abs/2105.14781 (2021)
29. Klein, T., Nabi, M.: Attention is (not) all you need for commonsense reasoning. CoRR, abs/1905.13497 (2019)
30. Geva, M., Gupta, A., Berant, J.: Injecting numerical reasoning skills into language models. CoRR, abs/2004.04487 (2020)
31. Ross Quinlan, J.: Induction of decision trees. Mach. Learn. **1**, 81–106 (1986)
32. Moro, S., Cortez, P., Rita, P.: A data-driven approach to predict the success of bank telemarketing. Decis. Support Syst. **62**, 22–31 (2014)
33. Wang, T., Rudin, C., Doshi-Velez, F., Liu, Y., Klampfl, E., MacNeille, P.: A Bayesian framework for learning rule sets for interpretable classification. J. Mach. Learn. Res. **18**(1), 2357–2393 (2017)

Approximability of Semigroups with Finiteness Conditions

Svetlana Korabelshchikova[1], Larisa Zyablitseva[1], Boris Melnikov[2(✉)], and Dang Van Vinh[3]

[1] Northern (Arctic) Federal University named after M.V. Lomonosov, Arkhangelsk, Russia
[2] Shenzhen MSU – BIT University, 1, International University Park Road, Dayun New Town, Longgang, Shenzhen 518172, China
bormel@mail.ru
[3] Ho Chi Minh City University of Technology, Ho Chi Minh, Vietnam

Abstract. This paper considers the concept of approximability of a semigroup with respect to a predicate and prepares the necessary basis for reducing the problem of isomorphism of semigroups to the problem of computing graph invariants. Two variants of mappings that map a semigroup to some graph are proposed. #COMESYSO1120.

Keywords: Semigroup · Graph · Semilattice

1 Representation of Finite Semigroups by Graphs

Consider one variant of mapping an oriented graph G_A to a semigroup A. The vertices of graph G_A will be given by elements of semigroup A. In this case an arc e between a, $c \in A$ is created only if $\exists b \in A$: $ab = c$.

If the correspondence is required to be one-to-one, then the arc $e = (a, c)$ must have weight b, but this is not mandatory. Here we will only consider the multiplicity of the arcs, but the arcs will be unweighted.

Example 1. Let A be a finite semigroup given by the Cayley table:

*	a	b	c	d
a	a	a	a	a
b	a	b	a	a
c	a	a	c	a
d	a	a	a	d

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 278–285, 2024.
https://doi.org/10.1007/978-3-031-53552-9_25

Then it will be matched with the orgraph G_A given by the adjacency matrix:

$$G = \begin{pmatrix} 4\,0 & 0\,0 \\ 3\,1 & 0\,0 \\ 3\,0 & 1\,0 \\ 3\,0 & 0\,1 \end{pmatrix}.$$

For some special kinds of semigroups, however, other correspondences can be considered. In particular, a commutative regular semigroup can naturally be matched to a semilattice of idempotents. If a commutative semigroup is idempotent, then this correspondence is mutually unambiguous up to isomorphism. Therefore semigroups S in which for arbitrary $x, y \in S$ the identities are true:

$$x^2 = x \text{ and } xy = yx$$

are also called "semilattice". Below we show two ways of matching a graph for such a semigroup.

Example 2. Let a semigroup $S = \{1, 2, \dots, n\}$ be given, and the operation between elements of the semigroup is defined by the following rule:

$$(\forall a, b \in S) a \cdot b = a \Leftrightarrow \text{a} \ \leq \ \text{b}.$$

For each element $a \in S$, $a \bullet a = a$ holds because $a \leq a$. It is also easy to see that S is a commutative semigroup. The semigroup S will be matched with the graph shown in Fig. 1. The graph of this semilattice and the semigroup S itself is called an n-element chain.

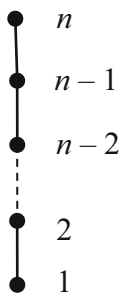


Fig. 1. Graph G_S of a semilattice S.

Example 3. Given a semigroup $A = \{z, 1, 2, \dots, n\}$, each element of which is an idempotent and the product of any two different elements is z - zero of semigroup A:

$$(\forall i, j \in A) i \cdot j = z \text{ if } i \neq j; \ (\forall i \in A) i \cdot i = i.$$

This semigroup corresponds to the graph shown in Fig. 2.

The represented graphs can be considered both undirected and oriented. One can specify an orientation of edges (for example, from top to bottom) and consider an oriented graph P_A and an undirected graph G_A corresponding to the initial semigroup A.

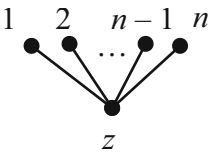


Fig. 2. Graph G_A of semilattice A.

Example 4. For semigroup A from Example 1 the graphs P_A and G_A have adjacency matrices:

$$P = \begin{pmatrix} 0\,0 & 0\,0 \\ 1\,0 & 0\,0 \\ 1\,0 & 0\,0 \\ 1\,0 & 0\,0 \end{pmatrix} \text{ and } G = \begin{pmatrix} 0\,1 & 1\,1 \\ 1\,0 & 0\,0 \\ 1\,0 & 0\,0 \\ 1\,0 & 0\,0 \end{pmatrix}.$$

In Examples 2, 3 and 4 the graphs corresponding to a semigroup are trees, but this is not mandatory. Quite often the graph of a semilattice contains a cycle. This is the case when for a pair of incomparable idempotents a and b there exists an exact upper edge in addition to the exact lower one. In this case a cycle.

$a \to \inf(a, b) \to b \to \sup(a, b) \to a$ occurs in graph G.

Example 5. Let us give examples of semilattices that are not trees (Fig. 3) and illustrate the inverse process of matching a graph with a semigroup given by the Cayley table.

Given commutativity, we do not fill the Cayley table for the semigroup A_1 under the main diagonal:

*	1	2	3	4	5	6	7	8	9	10	0
1	1	0	0	1	0	0	1	1	0	1	0
2		2	0	2	2	0	2	2	2	2	0
3			3	0	0	3	0	0	3	3	0
4				4	2	0	4	4	2	4	0
5					5	0	2	2	5	5	0
6						6	0	0	6	6	0
7							7	4	2	7	0
8								8	2	8	0
9									9	9	0
10										10	0
0											0

The Cayley table for semigroup A_2 is filled out as follows:

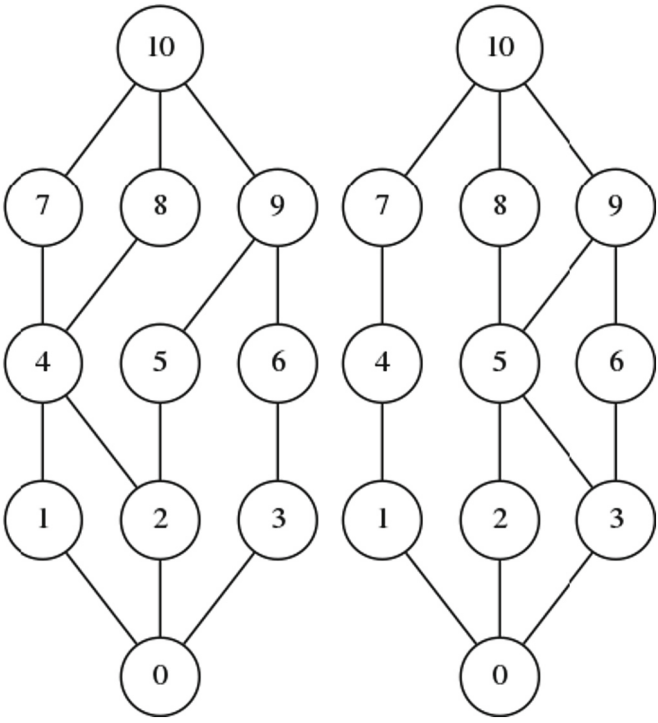


Fig. 3. Semilattices G_{A_1} and G_{A_2}.

*	1	2	3	4	5	6	7	8	9	10	0
1	1	0	0	1	0	0	1	0	0	1	0
2		2	0	2	0	0	2	2	2	2	0
3			3	0	3	3	0	3	3	3	0
4				4	0	0	4	0	0	4	0
5					5	3	0	5	5	5	0
6						6	0	3	6	6	0
7							7	0	0	7	0
8								8	5	8	0
9									9	9	0
10										10	0
0											0

The isomorphism of semilattices from Example 5 will be discussed later in the paper. Let us present the results concerning the ratio of semilattices and trees. For semilattices of small orders n these values were obtained in [1] (see Table 1).

Table 1. Ratio of semilattices to trees.

n	1	2	3	4	5	6	7	8
Number of semilattices	1	1	2	5	15	53	222	1078
Number of trees	1	1	2	4	9	20	48	115

Also [1] describes necessary and sufficient conditions that graphs of semilattice do not contain cycles and at the same time represent a connected graph.

Suppose that on the set M there is a binary partial-order relation denoted by $\leq$, and let $a \in M$. The set on which the linear order relation is given will be abbreviated as LOS. The set $L_a = \{x \in M \mid x \leq a\}$, consisting of all elements x of the set M satisfying the condition $x \leq a$, is the lower cone $NC(a)$ of the element a. Let a be an arbitrary element of the set under consideration, and for each such element $a \in M$, $NC(a)$ is a LOS. If the above conditions are satisfied, let us call the relation semi-linear from below.

Statement [1]. The representation of a commutative idempotent semigroup $(S, \cdot)$ by a graph G_S is a tree if and only if the natural partial order, which is defined by the relation

$$e \leq f \Leftrightarrow e \cdot f = f \cdot e = e,$$

is semilinear from below.

2 Finite Approximation of Semigroups

Mathematical science extensively uses the method of analogies, in which previously identified properties of some objects are transferred to other objects, in some characteristics similar to the original objects. For algebraic systems, such a method is the method of approximation. Algorithmization problems solvable for various objects, including graphs, can be considered from the perspective of abstract algebraic systems. This approach was widely used in the works of the Soviet mathematician A. I. Maltsev [2, 3] and his followers. In particular, algorithmic problems are inseparably connected with finite approximation of algebraic systems. Thus, in the theory of semigroups the approximation problems were dealt with by the "Leningrad algebraic school". (E. S. Lyapin, M. M. Lesokhin, I. S. Ponizovsky, S. I. Kublanovsky, and others) and the "Sverdlov school" (L.N. Shevrin, E.A. Golubov, M.V. Sapir, and others). A detailed review on approximation of semigroups and their applications is given in [4].

The finite approximability of a semigroup (group, ring, etc.) with respect to some predicate entails algorithmic solvability of the problem of occurrence of elements or subsets in the predicate in the algebraic system under consideration. Because this problem is so important, it has been well studied, and necessary and sufficient conditions for

the finite approximation of semigroups with respect to various predicates, such as predicates of an element to some special subset, subgroups with different properties, Green's relations, and some others, have been obtained. Historically, the first and most natural predicate was the equality predicate. Semigroups (groups) finitely approximated with respect to the equality predicate are called finitely approximated semigroups (groups).

Let us consider what other important conclusions follow from the finitively approximated approximation of initial semigroups (groups). The English mathematician Karl W. Gruenberg proved as early as the second half of the 1950s in [5] that "the free product of finitely approximated groups is finitely approximated. In the 1970s, however, E. A. Golubov [6] obtained a generalization of this fact to semigroups. If we take two FA semigroups and make their free product, the result will also be an FA semigroup.

At about the same time E.S. Lyapin [7] defined an essentially new notion, the notion of an independent product of two semigroups. It is, on the one hand, a combination of a free product of semigroups and groups, but, on the other hand, it no longer possesses the properties formulated above. As we managed to establish, the independent product of FA semigroups no longer has to be an FA semigroup at all, as the following example demonstrates.

Example 6. Consider a multiplicative semigroup with zero

$$S_1 = \{0, b, z_1, z_2, \ldots, z_n, \ldots, x_1, x_2, \ldots, x_n, \ldots\},$$

where multiplication is given according to the rule:

$$x_i x_j = \begin{cases} z_i, \text{ if } i = j, i = 1, 2, \ldots \\ b, \text{ if } i \neq j, j = 1, 2, \ldots \end{cases},$$

and all other pairwise multiplications are zero 0.

Also consider the second multiplicative semigroup with zero

$$S_2 = \{0, z_1, z_2, \ldots, z_n, \ldots, y_1, y_2, \ldots, y_n, \ldots\},$$

where the multiplication is given according to the rule:

$$y_i y_j = z_i, \text{ if } i = j, i = 1, 2, \ldots,$$

and all other pairwise products are zero 0.

It is easy to show that both semigroups S_1 and S_2 are FA semigroups. Let the semigroup S be an independent multiplication of the above semigroups S_1 and S_2. Let us prove that it is not an FA semigroup.

It is not difficult to see that the semigroup S is not finite. In particular, it has infinitely many elements of the form $x_i y_1$, where $i \in N$. Let φ be an arbitrary homomorphism of a semigroup S onto some finite semigroup A. Then find an element $a \in A$ that has at least two prototypes of the set $\{x_1 y_1, x_2 y_1, \ldots\}$ under the homomorphism φ.

Let $\varphi(x_i y_1) = \varphi(x_j y_1) = a$, $(i \neq j)$.

Further, given that $x_i x_i y_1 = 0$ and $x_i x_j y_1 = b y_1 \neq 0$, then elements 0 and $b y_1$ are different elements of semigroup S. Nevertheless, under homomorphism φ their images will coincide, because

$$\varphi(x_i x_i y_1) = \varphi(x_i)\varphi(x_i y_1) = \varphi(x_i)a$$
$$\varphi(x_i x_i y_1) = \varphi(x_i)\varphi\left(x_j y_1\right) = \varphi(x_i)a$$

This proves that the semigroup S is not an FA semigroup.

In [8] O.A. Snetkov revealed some special cases in which an independent product will be an FA semigroup if the initial semigroups are FA semigroups.

In particular, let S_1 and S_2 be FA semigroups whose intersection is a unit perfect semigroup. Consider their independent product S. Let us denote by

$$K = S_1 \cap S_2,$$

where K is a unit-ideal semigroup. Since by convention semigroups S_1 and S_2 are FA semigroups, we shall find homomorphisms φ and ψ of semigroups S_1 and S_2 respectively into some finite semigroup B, and $\varphi(k) = \psi(k)$ for all $k \in K$. In particular, we can take two copies of the same unit-perfect semigroup S_1, that is, $S_1 = S_2$

For each element x $\in S_1 \cup S_2$, compare the element $\chi(x)$ as follows:

$$\chi(x) = \begin{cases} \varphi(x), \text{ if } x \in S_1 \\ \psi(x), \text{ if } x \in S_2 \end{cases}.$$

Then the mapping χ: $\chi(x_1, x_2, .., x_n) = \chi(x_1)\chi(x_2) \ldots \chi(x_n)$ is a homomorphism of a semigroup S into a semigroup B and in some cases separates any two different elements of an independent product of S.

Let us describe these cases according to [8]; see also [9, 10].

Statement 2. If S is an FA semigroup and $s_1, s_2, .., s_n$ are pairwise distinct elements of it, then there exists a homomorphism φ of a semigroup S into a finite semigroup B such that $\varphi(s_1), \varphi(s_2), \ldots, \varphi(s_n)$ are pairwise distinct elements of B.

Statement 3. Let the semigroup S be an independent multiplication of semigroups S_1 and S_2 having unitwise perfect intersection $K = S_1 \cap S_2 = \{0\}$ - the zero of these semigroups. For S to be an FA semigroup, it is necessary and sufficient that S_1 and S_2 are FA semigroups.

Statement 4. Let the semigroups S_1 and S_2 be commutative bundles. Let S be their independent multiplication. Then the semigroup S is an FA semigroup.

Theorem. Let the semigroups S_1 and S_2 be Archimedean finitely generated semigroups in which the commutative condition is satisfied. Let $K = S_1 \cap S_2 = G$ be the maximal subgroup of semigroups S_1 and S_2 and S be their independent multiplication. For S to be an FA semigroup it is necessary and sufficient that S_1 and S_2 are FA semigroups.

3 Conclusion

Thus, this paper has proposed two variants of mappings that map a semigroup to a graph. At the same time the main results concerning the problem of decidability of the predicate by finite approximation are presented. Next, the isomorphism predicate given on the set of all semigroups will be considered. We will also apply the approximation method to the problem of establishing isomorphism of semigroups, and consider the composition of two mappings - the first mapping maps a semigroup to a graph, the second one computes the invariant of the graph. In the case where the obtained values are not equal, the conclusion is made that the semigroups are non-isomorphic. If the

range of values of the invariant is a finite set and the semigroups are approximated by this invariant with respect to isomorphism, we can say that the problem of isomorphism of semigroups for the class of semigroups under consideration is solvable.

References

1. Zyablitseva, L.V., Pestov, S.A.: Checking algorithm of semilattices isomorphism with the use of invariants of graph theory. Arct. Environ. Res. **17**(4), 368–375 (2017). https://doi.org/10.17238/issn2541-8416.2017.17.4.368
2. Maltsev, A.I.: Algebraic Systems. Fizmatlit, Moscow (1970)
3. Maltsev, A.I.: Algorithms and Recursive Functions. Nauka, Moscow (1986)
4. Lesokhin, M.M., Tutygin, A.G., Yashina, E.Y.: Approximation of semigroups. Partitioning and homomorphic representations of semigroups. A.I. Gertsen Russian State Pedagogical University, St. Petersburg, Russia (1992)
5. Gruenberg, K.W.: Residual properties of infinite soluble groups. Proc. London Math. Soc. **3**(25), 29–62 (1957)
6. Golubov, E.A.: Free product and entanglement of finitively approximated semigroups. Mat. zapas Ural'skii univ. **8**(1), 3–15 (1971)
7. Lyapin, E.S.: Intersection of independent subgroups of a semigroup. Izvestia Vuzov. Math. **4**, 67–73 (1970)
8. Snetkov, O.A.: Approximation of an independent product of semigroups. Bull. Izhevsk State Tech. Univ. **1**, 95–96 (2008)
9. Korabelshchikova, S.Yu., Melnikov, B.F., Melnikova, E.A., Zyablitseva, L.V.: Some approaches for determining isomorphism of semigroups of small order. IOP Conf. Ser. Mater. Sci. Eng., 52047 (2020). Krasnoyarsk Science and Technology City Hall of the Russian Union of Scientific and Engineering Associations
10. Zyablitseva, L.V., Korabelshchikova, S.Yu., Danilova T.V.: Representation of semigroups of right zeros and their semilattices by a transformation semigroup. News of Higher Educational Institutions. Volga Region. Phys. Math. Sci. **1**(61), 33–44 (2022)

Semigroup Invariants of Graphs with Respect to Their Approximability

Svetlana Korabelshchikova[1], Larisa Zyablitseva[1], Boris Melnikov[2(✉)], and Dang Van Vinh[3]

[1] Northern (Arctic) Federal University named after M.V. Lomonosov, Arkhangelsk, Russia
[2] Shenzhen MSU – BIT University, 1, International University Park Road, Dayun New Town, Longgang, Shenzhen 518172, China
bormel@smbu.edu.cn
[3] Ho Chi Minh City University of Technology, Ho Chi Minh, Vietnam

Abstract. This paper continues the consideration of the notion of approximability of a semigroup with respect to a predicate. This problem related to isomorphism of semigroups can be reduced to the problem of computing graph invariants. The paper investigates the problem of computing invariants in the case of representation of a finite graph by a adjacency matrix, and selects the invariants that most effectively distinguish non-isomorphic semigroups. A theorem has been proved stating that the diameter of a graph, the number of edges and the determinant of the adjacency matrix are not semigroup invariants for semigroups of a given order n with respect to the composition mapping. #COMESYSO1120.

Keywords: Semigroup · Graph · Isomorphism · Predicate

1 Introduction. The Problem of Isomorphism of Graphs and Semigroups

This paper is a continuation of the research reported in [1]. We continue to consider the notion of approximability of a semigroup with respect to a predicate. In this paper we prove a theorem stating that the diameter of a graph, the number of edges and the determinant of the adjacency matrix are not semigroup invariants for semigroups of a given order n with respect to the first mapping.

First of all, let us note three factors that attract attention to the problem of graph isomorphism. Firstly, this is an unsolved fundamental theoretical problem. Secondly, even its theoretical complexity has not yet been classified. According to the survey most of the scientists working in this field believe in the existence of polynomial algorithm for its solution. Thirdly, the apparatus of graph theory is truly universal and is used to solve a huge number of both applied and theoretical problems. Graph models are applied in many areas in modern research: in applied tasks of network planning [2, 3], in pattern recognition tasks [4], in models of parallel programming [5], in models required for creating heuristic algorithms [6, 7] and many others. It is especially important that mathematical models of social networks and some other models related to work in the

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 286–292, 2024.
https://doi.org/10.1007/978-3-031-53552-9_26

Internet are described in graph theory language [8, 9]. Therefore, the solution of graph isomorphism problem even for some particular cases is essential. Let us note that the very notion of graph isomorphism is formulated in different ways for different kinds of graphs: oriented and undirected, simple graphs and pseudographs. However the general meaning of the notion of isomorphism is as follows: is it possible to establish a bijection between vertices of two graphs such that these graphs coincide?

In this paper different variants of matching of graphs to a semigroup are considered and as a result algorithms for computing some graph invariants for differentiation of non-isomorphic semigroups can be used.

Let A be a set of semigroups and B a set of graphs. A mapping φ: $A \to B$ will be called approximative if for any $A_1, A_2 \in A$ it follows that $\varphi(A_1) \ncong \varphi(A_2)$ implies that $A_1 \ncong A_2$. On the basis of this correspondence, a solution to the problem of isomorphism of semigroups will be proposed and the effectively computable semigroup invariants of graphs will be distinguished [10]. To compute a number of invariants we will use the representation of a graph by a vertex adjacency matrix.

2 Semigroup Invariants of Graphs

Now consider the problem of isomorphism of semigroups. Consider an isomorphism predicate given on the set of all semigroups. Let us apply an approximation method to the problem of establishing isomorphism of semigroups. We compose a composition of two mappings - the first mapping maps a semigroup to a graph, the second mapping computes an invariant of the graph. If the obtained values are not equal, we conclude that the semigroups in question are non-isomorphic. If the range of values of the invariant is a finite set and the semigroups are approximated by this invariant with respect to isomorphism, we can say that the problem of isomorphism of semigroups for the class of semigroups under consideration is solvable.

The first mapping of φ_1. Let A be an arbitrary semigroup, that is, the only result of multiplication of any two elements is specified, and the given operation is associative. The elements of a semigroup A are vertices of a graph, and an edge between two different $a, c \in A$ occurs if $(\exists b \in A)\ ab = c$.

Let f_1, f_2, f_3, f_4, f_5 be the Wiener index, diameter, number of edges, determinant of the adjacency matrix, and ordered vector of degrees of vertices, respectively. Obviously, if for two semigroups A_1 and A_2 in at least one case out of five we obtain

$$f_i^{\circ}\ \varphi_1(A_1)\ \neq f_i^{\circ}\ \varphi_1(A_2),$$

then we can conclude that the semigroups are non-isomorphic. Thus the semigroups can be compared through computation of graph invariants. Graphs in this case are oriented, with multiple edges, and the mapping in question is approximative.

It is clear from the definition that the coincidence of graph invariants is a necessary condition for isomorphism of semigroups. This problem has also been studied in [11, 12].

The formulas for calculating the efficiency of invariants are given below.

$$P = \frac{C_{M_k}^2}{C_N^2} = \frac{M_k(M_k - 1)}{N(N - 1)}, \tag{1}$$

where N is the total number of semigroups of some order, M_k is the number of semigroups with a given value of the invariant.

The required probability is calculated as follows:

$$P = \sum_k \frac{M_k(M_k - 1)}{N(N-1)} = \frac{\sum_k M_k(M_k - 1)}{N(N-1)}. \tag{2}$$

The study was conducted for semigroups of small orders, the list of which can be extracted from the computer algebra system of GAP. We considered such invariants of graphs as the Wiener index W(G), the diameter of the graph d(G), the number of edges of the graph m(G), the determinant of the adjacency matrix det ⟦(G)⟧), the ordered vector of degrees of vertices v(G). The results of calculations are given in Table 1.

Table 1. Probabilities of matching invariants for the first variant of matching.

Number of elements n	2	3	4	5	6	7
Number of semigroups N	4	18	126	1160	15973	836021
Probability of matching $W(G)$	0.16667	0.17647	0.14108	0.11166	0.09966	0.16892
Probability of matching $d(G)$	0.50000	0.88889	0.98413	0.99828	0.99988	0.99999
Probability of matching $m(G)$	1	1	1	1	1	1
Probability of matching $det(G)$	0.16667	0.32026	0.43771	0.48644	0.58541	0.87271
Probability of matching $v(G)$	0.16667	0.05882	0.01918	0.00492	0.00280	0.01769

The smaller the value of the probability in Table 1, the more effectively the invariant distinguishes non-isomorphic semigroups. For this reason we introduce the following definition.

Definition 1. We will say that an invariant f recognizes non-isomorphic semigroups from class K if there exists a mapping φ from set K to graph set such that for any semigroup A_1 and A_2 of K such that $A_1 \ncong A_2$ it is true that $f \circ \varphi(A_1) \neq f \circ \varphi(A_2)$.

Any complete invariant recognizes non-isomorphic semigroups. However, the computation of complete invariants is known to have asymptotic complexity of the order of $O(n!)$, which is equivalent to solving the problem by the brute-force method. In the case where exact methods are not applicable, as in the traveling salesman's problem, we propose to use approximate methods that work in polynomial time. In this connection, we introduce the following definitions.

Definition 2. We will say that an invariant f recognizes non-isomorphic semigroups of class K with probability p with respect to mapping φ, if for any semigroup A_1 and A_2 of K, it is true that $f \circ \varphi(A_1) \neq f \circ \varphi(A_2)$ with probability p.

Definition 3. Invariants of graphs recognizing semigroups from class K with probability at least 0.75 with respect to mapping φ will be called semigroup invariants.

Consider such polynomially computable invariants of graphs as Wiener index $W(G)$, diameter of graph $d(G)$, number of edges of graph $m(G)$, determinant of adjacency matrix $\det(G)$, ordered vector of degrees of vertices $v(G)$.

Theorem 1. Let n be an arbitrary natural number, $n \geq 3$. Diameter of graph, number of edges and determinant of adjacency matrix are not semigroup invariants for semigroups of given order n with respect to the first mapping.

Proof. Let $d(G_A)$ be the diameter of the graph G_A, corresponding to semigroup A. Let us prove that it can take only two values. Firstly, prove that it cannot be greater than 1. By the contrary method, suppose that the diameter $\mathrm{d}(G_A) \geq 2$.

This means that there are three vertices x, y and z such that for the graph corresponding to the semigroup there exists an edge (x, y) and an edge (y, z) but no edge (x, z) or (z, x). We obtain:

$$\exists a', b' \in A : x \cdot a' = y \wedge yb' = z. \tag{3}$$

Multiplication in a semigroup is associative, so:

$$z = y \cdot b' = (x \cdot a') \cdot b' = x \cdot (a' \cdot b'), \text{ то есть } z = x \cdot (a' \cdot b'). \tag{4}$$

This means that there exists an edge (z, x).

Let us find out what values it can take. Let A be a semigroup of left-hand zeros. Then $\mathrm{d}(G_A) = 0$, since the condition in the semigroup $(A, \cdot)$ is satisfied:

$$\forall x \in A : \forall y \in A : x \cdot y = x. \tag{5}$$

Let n be the given order of the semigroup. Condition (5) in our case means that the corresponding graph has $n - 1$ loops at each vertex and no other edges. This means that the diameter of the semigroup is 0. Among all semigroups of order n only one satisfies relation (5), i.e. it is a left-zero semigroup. For it the diameter value will be 0, and for all other semigroups of a given order it will be 1.

Let $m(G_A)$ be the number of edges of graph G_A corresponding to semigroup A. Since in the first correspondence the multiplicity of edges is taken into account, the number of edges of a graph G_A corresponding to a semigroup A of order n is n^2, i.e. the same for all semigroups of given order. Obviously, this invariant is not a semigroup invariant since it coincides with any two semigroups of a given order at the first correspondence.

Now consider the invariant $\det(G_A)$, which is the determinant of the adjacency matrix of the graph G_A corresponding to the semigroup A (Table 2).

Table 2. Values of the determinant of the adjacency matrix of graph G_A.

Semigroup order	2	3	4	5	6	7
Probability of coincidence $\det(G)$	0.16667	0.32026	0.43771	0.48644	0.58541	0.87271

A study of the behavior of this invariant depending on the parameter n is presented in [9]; see also the references cited therein. As the order of a semigroup increases, the number of adjacency matrices whose determinant is 0 increases dramatically, which means that the probability of coincidence of the determinant values increases. Thus, for

semigroups whose order is greater than 2, the determinant of the adjacency matrix is not semigroup invariant with respect to the first mapping.

Thus we considered the first version of the correspondence ([1], Example 1) and the following polynomially computable invariants of a graph: number of edges, ordered vector of degrees of vertices, graph diameter, Wiener index and adjacency matrix determinant. Only the ordered vector of vertex degrees and the Wiener index are semigroup invariants for semigroups up to and including the 7th order.

Now let us consider the second variant of mapping a graph to a semigroup (item 1 of this paper, Examples 2–4). Let A belong to the class of semigroups. Let us map a semigroup A to an oriented graph P_A of partially ordered set $(A, \leq)$. The vertices of the graph P_A are also elements of the semigroup A, but the arc (b, a) between a pair of elements means that $a < b$ and there is no element c such that $a < c$ and $c < b$. One can also cancel the orientation of edges and consider an undirected graph G_A corresponding to semigroup A.

Vertices of the graph P_A having semidegree of the arrival 0 (maximal elements in the partially ordered set $(A, \leq)$) will be called end vertices. A vertex of graph P_A having semidegree of the outcome 0 (the minimal element in the partially ordered set $(A, \leq)$) will be called a root.

If graphs P_A and G_A do not contain cycles, then effective algorithms for checking isomorphism of trees can be applied to them. Otherwise we can use the approach based on the computation of invariants.

Note that the second variant of matching graphs to a semigroup is not suitable for all semigroups. One can apply it to commutative regular semigroups by matching a semigroup to a semigroup with the semilattice of its idempotents. However, this approach will map a large number of different semigroups to the same semilattice. This means that the efficiency of any invariant in recognizing semigroups in such a matching will be low.

If we consider the second version of matching only on commutative idempotent semigroups, we can specify semigroup invariants for semigroups of small orders. Consider the following set of invariants of a semilattice, which afterwards we will call invariant 1. For each vertex we will match a vector consisting of three components:

- an ordered vector of semidegrees of the outcomes of vertices in the graph P_A;
- an ordered vector of semidegrees of vertices in the graph P_A;
- a vector of distances from the given vertex of the graph to the endpoints and the root in the undirected graph G_A.

All obtained vectors are ordered lexicographically and the vector composed in this order is called invariant 1. Note that invariant 1 distinguishes semilattices from Example 5 of this paper. This means that the semilattices A_1 and A_2 are non-isomorphic. Denote by $N(i)$ the maximum number of semilattices of order i with coincident invariant 1. The results of calculations are given in Table 3.

The results in Table 3 show that the invariant 1 is a complete semigroup invariant for all commutative idempotent semigroups up to and including order 7.

Table 3. Maximum number of semilattices with matching invariant 1.

Semigroup order	1	2	3	4	5	6	7	8
$N(i)$	1	1	1	1	1	1	1	2

3 Main Theorem

As the result of the research two approximating mappings are distinguished from the considered different variants of mappings of graphs to semigroups. The first mapping is defined on the set of all finite semigroups, the second one on the set of commutative idempotent semigroups (semilattices), possibly infinite. Semigroup invariants, i.e. invariants having efficiency at least 0.75 for semigroups of a given order, are distinguished from polynomially computable graph invariants. The following general theorem, valid for all natural $n \geq 3$, is proved.

Theorem 2. Let n be an arbitrary natural number, $n \geq 3$. The diameter of the graph, the number of edges and the determinant of the adjacency matrix are not semigroup invariants for semigroups of a given order n with respect to the first mapping.

This theorem is a generalization of the results obtained in [11] for the case of arbitrary natural N. Algebraic methods and their results for graphs and semigroups were published by one of the authors in [11, 12]. As a result of this research, the invariants that are semigroup invariants are distinguished. A complete invariant, which has an efficiency of 1 and is a complete semigroup invariant, was found for the second correspondence. This is a set of three invariants of graph at the second correspondence, which is an ordered vector of length $3n + mn$, where n is the number of graph vertices, m is the number of end vertices. This vector (invariant 1) contains information about semidegrees of arrival and outcome of vertices in the corresponding orgraph, and also information about distances from each vertex to the end vertices and the root in the corresponding undirected graph.

4 Conclusion

The paper proposes two variants of approximation mappings from the set of semigroups to the set of graphs. The first mapping is defined on the set of all finite semigroups, the second one on the set of commutative idempotent semigroups (semilattices). The composition of such a mapping with the graph invariant function in some cases allows to distinguish non-isomorphic semigroups.

In this paper we investigated the problem of calculating invariants in the case of representation of a finite graph by a adjacency matrix, and carried out the selection of invariants that most effectively distinguish non-isomorphic semigroups (semigroup invariants).

The method of recognition of non-isomorphic semigroups applied in this paper, of course, cannot completely solve the problem of isomorphism of semigroups as well as the problem of isomorphism of graphs corresponding to them. However, it shows good results for semigroups and graphs of small order. We believe that a promising direction

for further research is the application of the proposed methods and algorithms for graphs and semigroups of larger orders.

References

1. Korabelshchikova, S., Zyablitseva, L., Melnikov, B., Vinh, D.V.: Approximability of semigroups with finiteness conditions. Lecture Notes in Networks and Systems (2023)
2. Zhang, X., Moore, C., Newman, M.E.J.: Random graph models for dynamic networks. Eur. Phys. J. B **90**(10), 200 (2017)
3. Bulynin, A., Meshchanin, V., Melnikov, B., Terentyeva, J.: Algorithms for designing communication networks using greedy heuristics of various types. In: CEUR Workshop Proceedings, pp. 273–276 (2020). Art. 2667
4. Poddar, A., Sahidullah, M., Saha, G.: Speaker verification with short utterances: a review of challenges, trends and opportunities. IET Biometrics **7**(2), 91–101 (2018). https://doi.org/10.1049/iet-bmt.2017.0065
5. Melnikov, B.F., Panin, A.G.: Parallel implementation of multieuristic approach in the problem of comparison of genetic sequences. Sci. Vector Togliatti State Univ. **4**(22), 83–86 (2012)
6. Melnikov, B.F., Sciarini-Guryanova, N.V.: Possible edges of a finite automaton defining a given regular language. Korean J. Comput. Appl. Math. **9**(2), 475–485 (2002)
7. Hromkovic, J., Staub, J.: The problem with debugging in current block-based programming environments. Bull. EATCS **135**, 667 (2021)
8. Raigorodskii, A.M.: Small subgraphs in preferential attachment networks. Opt. Lett. **11**(2), 249–257 (2015). https://doi.org/10.1007/s11590-015-0945-9
9. Melnikov, B.F., Sayfullina, E.F., Terentyeva, Y.Y., Churikova, N.P.: Application of algorithms for generating random graphs for investigating the reliability of communication networks. Inf. Commun. **1**, 71–80 (2018)
10. Zyablitseva, L.V., Pestov, S.A.: Checking algorithm of semilattices isomorphism with the use of invariants of graph theory. Arct. Environ. Res. **17**(4), 368–375 (2017). https://doi.org/10.17238/issn2541-8416.2017.17.4.368
11. Zyablitseva, L.V.: Efficiency of invariants for semigroups of small orders. In: Intellectual and Information Technologies and Intelligent Business (INFOS 2019), pp. 24–27 (2019)
12. Zyablitseva, L.V., Pestov, S.A.: Test algorithms of the graph isomorphism checking in the theory of semigroup. Arct. Environ. Res. Ser. Nat. Sci. **4**, 69–74 (2016). https://doi.org/10.17238/issn2227-6572.2016.4.69

Algorithmizing Aspects of Some Combinatorial Block-Designs Implemented in Network Systems

Alexander Frolov(✉), Natalya Kochetova, and Anton Klyagin

National Research University Moscow Power Engineering Institute, Moscow, Russian Federation
{frolovab,kochetovanp,kliaginao}@mpei.ru

Abstract. The main numerical characteristics of a number of combinatorial block designs (CBD), which are essential for choosing parameters of network systems of combinatorial block structure are given. In addition to the traditionally considered blocks of CBD, attention is paid to dual blocks, formed by combinations of blocks containing a certain element. By combining blocks and dual blocks in one block design the so-called combined CBD is generated, the elements of which are pairs (block, element of this block). This achieves the scaling effect of the network system. The stage of its synthesis consists in the selection of CBD in accordance with the requirements of the parameters of the network system and its subsystems, as well as the links between subsystems. The main result are sets of algorithms for computing blocks, dual blocks as well as intersections of two blocks and two dual blocks for systems of Steiner triples, projective and affine geometries and transversal combinatorial block-designs. They are used to derive algorithms for computing lists of blocks, each of which intersects each of two given non-intersecting blocks, as well as algorithms for constructing block designs in general. The algorithms are given in two notations - algebraic and numeric, i.e. with representation of inputs and outputs by identifiers of elements or blocks or by numbers of these identifiers at a certain numbering by initial non-negative integers. For this purpose, the corresponding numbering algorithms and numbering inversion algorithms are used. #COMESYSO1120

Keywords: combinatorial block design · combinatorial network structure · networks scaling · dual blocks · algorithms · blocks intersection · dual blocks intersection

1 Introduction

Combinatorial block designs (CBD) [1, 2] are used as mathematical models of network systems such as computer networks [3, 4] and key systems in wireless sensor networks (WSN) [1, 5], in particular on the standard [6] and ZigBee technology [7–9]. When combining both blocks and dual blocks in one system, CBD acquires the structure of a combined combinatorial design (CCBD), and its scaling effect appears [10, 11].

The task of synthesizing network structures of the class under consideration is to select suitable CBDs and their characteristics, the latter must correspond to the technical

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 293–329, 2024.
https://doi.org/10.1007/978-3-031-53552-9_27

parameters of the system being synthesized, such as the scale (number of nodes), the number of nodes in subsystems, the length of the shortest route between nodes, etc.

Next, the problem arises of calculating the composition of subsystems, as well as the connections between nodes. This corresponds to the problems of constructing blocks and dual blocks of CBD or CCBD chosen at the stage of synthesis, as well as their intersections. Until now, the algorithmization of the CBD has been limited to the problems of constructing CBD blocks and their intersections [3, 12–15] and no attention has been paid to dual blocks. The exception is [16], where the dual blocks of affine geometry and their intersections are used, but the blocks are not explicitly used. In a number of works, algebraic identifiers of elements and blocks are used to solve these problems. Note that different types of CBDs can be used in the same network system [13, 14].

In this paper, we propose a general method for the CBD and CCBD algorithmization and provide algorithms for constructing blocks and dual blocks using this method, as well as the intersections of two blocks and two dual blocks, and in its absence, determining the block which with each of two nonintersecting blocks intersects, as applied to projective planes PP(2,n), corresponding residual block designs such as affine planes RPP(2,n), linear transversal combinatorial block designs TD(2,k,n), as well as the Steiner triple systems STS($6t + 3$) and STS($6t + 1$). The relevance of the problem being solved is due to the need for these algorithms in the construction of real systems of a combinatorial-block structure. The methods of this paper are also applicable to create similar sets of algorithms for projective geometries PG(3,n) and affine geometries RPG(3,n) as well as quadratic combinatorial block designs TD(3,k,n).

2 Methods

2.1 Parameters of Network Systems with Combinatorial Block Design Structure

CBD is a system of sets (**X,B**), where **X** is a set of elements, **B** is a set of blocks B, $B \in \mathbf{B}$, $|\mathbf{X}|=v$, $|B|=k$. We also define the set **D** of dual blocks $D(e)$ as sets of blocks containing the element e: $D(e)=\{B:e \in B\}$, $|D(e)|=r$. It is clear that $|\mathbf{D}|=v$. CBD (**X,B**) corresponds to the combined combinatorial block design (CCBD) (**M,X,B,D**), where **M** is the set of elements of CCBD that are pairs $<B,e>, e \in B$, or, equivalently, $<B, D(e)>$, $e \in B$ [10]. The set of pairs with the same first elements constitutes the CCBD block, and the set of pairs with the same second elements constitutes the dual CCBD block. Each element of the CCBD belongs to one block and one dual block of the CCBD. CBD and CCBD are mathematical models of network systems of the corresponding structure. In them, blocks and dual blocks correspond to subsystems, each node of which belongs to two subsystems of one (numbers are given in italics) and another (dashed links) (Fig. 1).

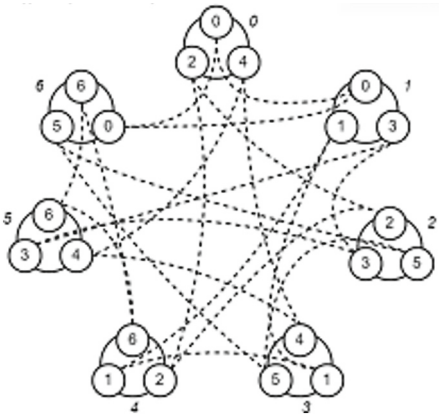


Fig. 1. A network system with the structure of a combined Steiner triple system of order 7.

A feature of the use of combinatorial block designs in the synthesis of network systems is fixed sets of numerical characteristics, determined by the parameters of the CBD and CCBD. In the example above, there are 21 nodes, 7 internal and 7 external subnets, each with three nodes. These numbers are defined by order 7 of the Steiner triple system. Currently, the projective planes PP(2,n) and projective geometries PG(3,n), affine planes RPG(2,n) and affine geometries RPG(3,n), transversal CBD: linear TD(2,k,n) and quadratic TD(3,k,n) combinatorial block-designs, Steiner triple systems STS($6t+3$) and STS($6t+1$), as well as unitals of order n are used as mathematical models of networks systems.

2.2 On Algorithmization of CBD and CCBD

To implement network systems with a CBD or SCBD structure, algorithms for constructing blocks and dual blocks are needed, and for routing in them or searching for common keys, algorithms for computing block intersections, as well as dual blocks, are also required.

The CBDs mentioned in the introduction are constructed in some algebraic system and are specified by a numerical parameter n, which is simultaneously a system parameter of the system being synthesized. In this system, algebraic identifiers are set for each element, block and dual block selected at the stage of synthesis, by which the block element or dual block is computed in this algebraic system. Let us denote by $\tilde{e}$, $\tilde{B}$ the algebraic identifiers of elements and blocks, $\widetilde{\mathbf{X}}$, $\widetilde{\mathbf{B}}$ are the sets of such identifiers. Clearly, the identifiers of dual blocks $D(e)$ are identifiers of corresponding elements e. The inputs of algorithms are identifiers of operands, the outputs are identifiers of results. We call such notation of algorithms as algebraic notation. When using them in practice, it is more convenient to use a numerical notation of algorithms, when the input data and the result are represented by numbers of their identifiers. For this purpose, algorithms in algebraic notation are supplemented by algorithms of numbering identifiers of elements

and blocks, applied to output data and algorithms of numbering of elements and blocks for construction of identifiers on their numbers on the input. It is clear that such algorithms are created for a specific type of CBD. A set of such algorithms for projective planes PP(2,*n*) is given in [17]; in its creation the system of algebraic identifiers proposed in [13] was used.

This paper presents sets of algorithms in algebraic notation, including algorithms

- construction of blocks $\tilde{B} \to \overset{\sim k}{\boldsymbol{X}}$ and dual blocks $\overset{\sim}{\boldsymbol{X}} \to \overset{\sim r}{\boldsymbol{B}}$,
- calculation of intersections of pairs of blocks $\overset{\sim 2}{\boldsymbol{B}} \to \overset{\sim *}{\mathbf{X}}$ and pairs of dual blocks $\overset{\sim 2}{\boldsymbol{X}} \to \overset{\sim *}{\mathbf{X}}$,
- definitions of blocks blocks intersecting with each of two non-intersecting blocks if there are such pairs of blocks in CBD: $\overset{\sim 2}{\boldsymbol{B}} \to \overset{\sim *}{\mathbf{B}}$, for CBD and CCBD of the following types:

- systems of Steiner triples of order *n*, $n = 6t + 3$, $t \geq ,1$
- systems of Steiner triples of order *n*, $n = 6t + 1$, $t \geq 1$,
- projective planes PP(2,n) and projective geometries PG(3,n) whose order n is a prime number p or a power of a prime number *pl*, $l \geq 1$,
- affine planes RPP(2,n) and affine geometries RPG(3,n), order *pl*, $l \geq 1$,
- linear transversal CBD of order $n = pl$, $l \geq 1$,
- quadratic transversal CBDs of order *pl*, $l \geq 1$.

It is assumed that the indicated numerical characteristics and the numbers derived from them are global and are not indicated among the formal parameters of the algorithms.

For the purposes of visualization of the representation of CBD and their blocks and dual blocks, as well as for the convenience of using the algorithms, algorithms for numbering elements and blocks for each of the CBDs considered in the article are given, which made it possible to obtain numerical notations of the above algorithms. The numerical representations of CBDs, their blocks, dual blocks, as well as the intersections of blocks and dual blocks given in final section, can be considered as a certain set of tests for these algorithms.

Thus, the general technique used in this work for constructing CBD components consists in choosing a system of identifiers for elements and blocks of CBD, using an algebraic dependence of an identifier at the input and output of the algorithm, and constructing an unselected algorithm for converting a set of input identifiers into a set of output identifiers. The algorithms for numbering and reversing the numbering of identifiers convert the indicated algebraic notations of the algorithms into numerical notations.

3 Results: Algorithms

3.1 Algorithms for Systems of Steiner Triples [1]

The Bose construction. Let $(X,\circ)$ be an idempotent quasi-group of order $v = 6t + 3$, $t \geq 1$ with the operation $x \circ y = (n + 1)/2\, x * y \bmod n$, $n = 2t + 1$. $\widetilde{\mathbf{X}}$ is the set of identifiers $[x,i]$, $x \in X$, $X = \{0,\ldots v - 1\}$, $i \in \{1,2,3\}$, of the Steiner triple system elements. They are also the identifiers of dual blocks. A block is a triple (e_1,e_2,e_3) of elements with identifiers $[x,i],[y,i],[x \circ y,(i + 1) \bmod 3]$. The identifier of this block is the triple $[x, y, i]$.

Table 1 gives algorithms for constructing blocks and dual blocks defined by their identifiers $[x, y, i]$ and $[x, i]$ [1]. In Table 2 there are represented algorithms for the intersection a pair of blocks specified by identifiers $[x_1,y_1,i_1]$, $[x_2,y_2,i_2]$ and for the intersection a pair of dual blocks specified by identifiers $[x_1, i_1]$ and $[x_2, i_2]$. Parameters t, n are global.

Table 1. Bose STS-block and STS-dual block building algorithms

Алгоритм **Block_STS_Bose**(*id*)	Алгоритм **Dual_Block_STS_Bose**(*id*)
Input: *id*= [*x, y, i*]. Output: list of block elements identifiers	Input: *id* = [*x*, *i*]. Output: list of identifiers of block, containing given element

if *x*=*y*:
  return [[*x*,0],[*x*,1],[*x*,2]]
else:
  return
  [[*x*,*i*],[*y*,*i*],[*x*○*y*,(*i*+1)%3]]
	D = [*x*, *x*, 0]
	for *v* in range(*x* + 1, *n*):*D*.append([*x, v, i*])
	for *v* in range(0, *x*):*D*.append([*v, x, i*])
	z = 2*x* % *n*
	for *v* in range(*x* + 1, *x* + *t* + 1):
	  z_vmod=(*z* - *v*) % *n*
	  if *v* % *n* < *z_vmod n*
	  D.append([*v* % *n*, *z_v mod*, (*i* + 2) % 3])
	  else:
	  D.append([(*z_vmod*, *v* % *n*, (*i* + 2) % 3])
	return *D*

The Skolem construction. Let $(X,\circ)$ be a semi-idempotent quasi-group of order $v = 6t + 1$, $t \geq 1$ $x \circ y = \pi(x + y)$, where

$$\pi(x) = \begin{cases} \frac{x}{2}\ if\ x\ is\ even, \\ \frac{x+n-1}{2}, if\ x\ is\ odd, n = 2t. \end{cases}$$

Table 2. Bose STS-blocks and STS-dual blocks intersection building algorithms

Алгоритм **Blocks_STS_Bose_Inersection(***id*1,*id*2**)**	Алгоритм **Dual_Blocks_STS_Bose_Inersection(***id*1, *id*2**)**
Input: $id1 = [x_1, y_1, i_1]$, $id2 = [x_2, y_2, i_2]$. Output: common element identifier or []	Input: $id1 = [x_1, i_1]$, $id2 = [x_2, i_2]$. Output: the identifier of block, containing a pair of given elements

if $x_1 = y_1$ and $x_2 = y_2$: return []
if $x_1 = y_1$ and ($x_1 = x_2$ or $x_1 = y_2$):
  return $[x_1, i_2]$
if $x_2 = y_2$ and ($x_2 = x_1$ or $x_2 = y_1$):
  return $[x_2, i_1]$
$z_2 = x_2 \circ y_2$
if $x_1 = y_1$and $x_1 = z_2$:
  return $[z_2, (i_2 + 1)\% 3]$
$z_1 = x_1 \circ x_2$
if $x_2 = y_2$ and $x_2 = z_1$:
  return $[z_1, (i_1 + 1)\% 3]$
if $i_1 = i_2$ and ($x_1 = x_2$ or $x_1 = y_2$):
  return $[x_1, i_1]$
if $i_1 = i_2$ and ($y_1 = x_2$ or $y_1 = y_2$) :
  return $[y_1, i_1]$
if $i_1 = i_2$ and $z_1 = z_2$:
  return $[z_1, (i_1 + 1)\% 3]$
if $i_1 = (i_2 + 1)\% 3$ and $x_1 = z_2$:
  return $[x_1, i_1]$
if $i_1 = (i_2 + 1)\% 3$ and $y_1 = z_2$:
  return $[y_1, i_1]$
if $(i_1 + 1)\% 3 = i_2$ and $x_2 = z_1$:
  return $[x_2, i_2]$
if $(i_1 + 1)\% 3 = i_2$ and $y_2 = z_1$:
  return $[y_2, i_2]$
return []
	if $x_1 = x_2$:
	  return $[x_1, x_1, 0]$
	if $i_1 = i_2$:
	  if $x_1 < x_2$:
	  return $[x_1, x_2, i_1]$
	  else:
	  return $[x_2, x_1, i_1]$
	if $i_2 = (i_1 + 2) \% 3$:
	  if $x_2 < (2x_1 - x_2)\%n$::
	  return $[x_2, (2x_1 - x_2) \% n, i_2]$
	  else:
	  return $[(2x_1 - x_2) \% n, x_2, i_2]$
	if $i_1 = (i_2 + 2) \% 3$:
	  if $x_1 < (2x_2 - x_1)\%n$:
	  return $[x_1, (2x_2 - x_1) \% n, i_1]$
	  else:
	  return $[(2x_2 - x_1) \% n, x_1, i_1]$

$\widetilde{\mathbf{X}}$ is the set of identifiers $[x,i]$, $x \in X$, $X = \{0,\ldots v-1\}$, $i \in \{1,2,3\}$ and '*infty*' elements of the Steiner triple system. A block is a triple (e_1,e_2,e_3) of elements with identifiers $[x,0],[x,1],[x,2];[x,i],[y,i],[x^{\circ}y,(i+1) \bmod 3]$ or ['*infty*'],$[x+t,i],[x,i+1 \bmod 3]$. The identifier of this block is the triple $[x,y,i]$. Table 3 gives algorithms for constructing blocks and dual blocks given by their identifiers $[x, y, i]$ and $[x, i]$, in Table 4 algorithms for intersection a pair of blocks specified by identifiers $[x_1, y_1, i_1]$, $[x_2, y_2, i_2]$ and intersection a pair of dual blocks specified by identifiers $[x_1, i_1]$ и $[x_2,i_2]$. Parameters *t, n* are global.

3.2 Algorithms for Projective Planes PP(2,*n*) Based on Block Bases

The set of elements is the set of points of the projective space of dimension three i.e. the set of all normalized polynomials of the field GF(n^3), $n = p^l$, p is a prime. The elements e identifier is a list of the coefficients of the corresponding polynomial. Let x be a primitive element of the field GF(n); $x^0 = 0$; $x^j \in$ GF(n), $j = 1,\ldots,n-1$. A block is a set of points that includes two non-collinear points e_1 and e_2 that form the basis of the block (see Table 5). The block identifier is also a normalized polynomial of the field GF(n^3), one-to-one corresponding to a pair of basis element identifiers which algebraic closure is a block. Block identifiers and elements of the projective space are numbered by the initial elements of the extended natural series.

Table 6 gives algorithms for constructing blocks and dual blocks given by their identifiers $[a, b, c]$ and $[x,y,z]$. Table 7 represents algorithms for the intersection a pair of blocks specified by identifiers $[a_i, b_i, c_i]$, $[a_j, b_j, c_j]$ and the intersection a pair of dual blocks specified by identifiers $[x_i, y_i, z_i]$, $[x_j, y_j, z_j]$. Parameters *p, l, n* are global.

3.3 Algorithms for Projective Geometries PG(3,*n*) Based on Block Bases

The set of elements is the set of points of the projective space of dimension four i.e. the set of all normalized polynomials of the field GF(n^4), $n = p^l$, p is a prime. The elements e identifier is a list of the coefficients of the corresponding polynomial. Let x be a primitive element of the field GF(n); $x^0 = 0$; $x^j \in$ GF(n), $j = 1,\ldots,n-1$. A block is a set of points that includes three non-collinear points e_1, e_2, and e_3 that form the basis of the block. The block identifier is also a normalized polynomial of the field GF(n^3), one-to-one corresponding to a triple of basis element identifiers which algebraic closure is a block. Block identifiers and elements of the projective space are numbered by the initial elements of the extended natural series (see Table 8).

Table 3. Skolem STS-block and STS-dual block building algorithms

Алгоритм Block_STS_Skolem(*id*)	Алгоритм Dual_Block_STS_Skolem(*id*)
Input: *id*= [*x, y, i*]. Output: list of block elements identifiers	Input: *id*. Output: list of identifiers of block, containing given element

if *x*=*y*:
  return [[*x*,0],[*x*,1],[*x*,2]]
else if *x* != '*infty*':
  return
  [[*x*,*i*],[*y*,*i*],[*x*○*y*,(*i*+1)%3]]
else:
  return
  ['*infty*',[(*y*+*t*),*i*],[*y*,(*i*+1)%3]]
	D = []
	if *id* != '*infty*':
	  [*x*, *i*] = *id*
	  if *x* < *t*:
	  D.append([*x*, *x*, 0])
	  for *y* in range(*x* + 1, *n*):
	  D.append([*x*, *y*, i])
	  for *y* in range(0, *x*):
	  D.append([*y*, *x*, *i*])
	  if *x* < *t*:
	  *z*1 = 2*x* % *n*
	  for *v* in range(*x* + 1, *x* + *t*):
	  w = (*z*1 - *v*) % *n*
	  if *w* < *v* % *n*:
	  D.append(([*w*,*v*%*n*,
	  (*i*+2)%3]))
	  else:
	  D.append(([*v*%*n*,*w*,
	  (*i*+2)%3]))
	  D.append((['*infty*',*x*,(*i*+2)%3]))
	  else:
	  *z*2 = (2*x* + 1)% *n*
	  for *v* in range(*x* + 1, *x* + *t* + 1):
	  w = (*z*2 - *v*) % *n*
	  if *w* < *v* % *n*:
	  D.append(([*w*,*v*%*n*,
	  (*i*+2)%3]))
	  else:
	  D.append(([*v*%*n*,*w*,
	  (*i*+2)%3]))
	  D.append(['*infty*', *x* - *t*, *i*])
	else:
	  for *y* in range(0, *n* // 2):
	  for *i* in range(0, 3):
	  D.append(['*infty*',*y*,(*i*+1)%3])
	return *D*

Table 4. Skolem STS-blocks and STS-dual blocks intersection building algorithms

Algorithm **Blocks_STS_Skolem_ Inersection(***id*1.*id*2)	Algorithm **Dual_Blocks_ STS_SkolemInersection(***id*1.*id*2**)**
Input: *id*1= $[x_1, y_1, i_1]$, *id*2= $[x_2, y_2, i_2]$. Output: common element identifier or []	Input: *id*1 =$[x_1,\ i_1]$, *id*2= $[x_2,\ i_2]$. Output: the identifier of block, containing a pair of given elements

if $x_1 = y_1$ and $x_2 = y_2$:
return []
if x_2 != '*infty*' and x_1 != '*infty*':
$z_1 = x_1 \circ y_1$
$z_2 = x_2 \circ y_2$
if $x_1 = y_1$ and ($x_1 = x_2$ or $x_1 = y_2$):
return $[x_1, i_2]$
if $x_2 = y_2$ and ($x_2 = x_1$ or $x_2 = y_1$):
return $[x_2, i_1]$
if $x_1 = y_1$ and $x_1 = z_2$:
return $[z_2, (i_2 + 1)\ \%\ 3]$
if $x_2 = y_2$ and $x_2 = z_1$:
return $[z_1, (i_1 + 1)\ \%\ 3]$
if $i_1 = i_2$ and ($x_1 = x_2$ or $x_1 = y_2$):
return $[x_1, i_1]$
if $i_1 = (i_2 + 1)\ \%\ 3$ and $x_1 = z_2$:
return $[x_1, i_1]$
if $i_1 = i_2$ and ($y_1 = x_2$ or $y_1 = y_2$):
return $[y_1, i_1]$
if $i_1 = (i_2 + 1)\ \%\ 3$ and $y_1 = z_2$:
return $[y_1, i_1]$
if $(i_1 + 1)\ \%\ 3 = i_2$ and $x_2 = z_1$:
return $[x_2, i_2]$
if $(i_1 + 1)\ \%\ 3 = i_2$ and $y_2 = z_1$:
return $[y_2, i_2]$
if $i_1 = i_2$ and $z_1 = z_2$:
return $[z_1, (i_1 + 1)\ \%\ 3]$
if x_1 != '*infty*' and x_2 = '*infty*':
$z_1 = x_1 \circ y_1$
if $x_1 = y_1$ and $x_1 = y_2$:
return $[x_1, (i_2 + 1)\ \%\ 3]$
if $x_1 = y_1$ and $x_1 = y_2 + t$:
return $[x_1, i_2]$
if $i_1 = i_2$ and $x_1 = y_2 + t$:
return $[x_1, i_1]$
if $i_1 = (i_2 + 1)\ \%\ 3$ and $x_1 = y_2$:
return $[x_1, i_1]$
if $i_1 = i_2$ and $y_1 = y_2 + t$:
	if (id_1 != '*infty*') and (id_2 != '*infty*'):
	$[x_1, i_1] = id_1$
	$[x_2, i_2] = id_2$
	if $x_1 = x_2$ and $x_1 < t$:
	return $[x_1, x_1, 0]$
	if $i_1 = i_2$:
	if $x_1 < x_2$: return $[x_1, x_2, i_1]$
	else: return $[x_2, x_1, i_1]$
	if $i_1 = (i_2 + 2)\ \%\ 3$:
	if $x_1 = (x_2 + t)\ \%\ n$ and $x_2 < x_1$:
	return ['*infty*', x_2, i_1]
	if $x_2 < t$: $z = 2x_2$
	if $x_2 >= t$: $z = (2x_2 + 1)\ \%\ n$
	if $x_1 < (z - x_1)\ \%\ n$:
	return $[x_1, (z - x_1)\ \%\ n, i_1]$
	else:
	return $[(z - x_1)\ \%\ n, x_1, i_1]$
	if $i_2 = (i_1 + 2)\ \%\ 3$:
	if $x_2 = (x_1 + t)\ \%\ n$ and $x_1 < x_2$:
	return ['*infty*', x_1, i_2]
	if $x_1 < t$: $z = 2x_1$
	if $x_1 >= t$: $z = (2x_1 + 1)\ \%\ n$
	if $x_2 < (z - x_2)\ \%\ n$:
	return $[x_2, (z - x_2)\ \%\ n, i_2]$
	else:
	return $[(z - x_2)\ \%\ n, x_2, i_2]$
	if (id_1!='*infty*') and (id_2='*infty*'):
	$[x_1, i_1] = id_1$
	if $x_1 < t$:
	return ['*infty*', x_1, $(i_1 + 2)\ \%\ 3$]
	else: return ['*infty*', $x_1 - t$, i_1]
	if (id_2 != '*infty*') and (id_1 = '*infty*'):
	$[x_2, i_2] = id_2$

(*continued*)

Table 4. (*continued*)

```
    return [y1, i1]
  if i1 = (i2 + 1) % 3 and y1 = y2:
    return [y1, i1]
  if z1 = y2 and i1 = i2:
    return [y2, (i2 + 1) % 3]
  if z1= y2 + t and (i1+1) % 3 = i2:
    return [y2 + t, i2]
if x2 != 'infty' and x1 = 'infty':
  z2 = x2∘y2
  if x2 = y2 and x2 = y1:
    return [x2, (ii + 1) % 3]
  if x2 = y2 and x2 = y1 + t:
    return [x2, i1]
  if i1 = i2 and x2 = y1 + t:
    return [x2, i2]
  if i2 = (i1 + 1) % 3 and x2 = y1:
    return [x2, i2]
  if i1 = i2 and y2 = y1 + t:
    return [y2, i2]
  if i2 = (i1 + 1) % 3 and y1 = y2:
    return [y2, i2]
  if z2 = y1 and i1 = i2:
    return [y1, (i1 + 1) % 3]
  if z2= y1 + t and (i2+1) % 3 = i1:
    return [y1 + t, i1]
if x1 = 'infty' and x2 = 'infty':
  return ['infty']
return []
```

```
if x2< t:
  return ['infty', x2,
          (i2 + 2) % 3]
else: return ['infty',x2 - t, i2]
```

Table 5. PP2 Block bases

Block (or element) number	t=0	t	$t=jn+i$, $j=\overline{1,n}$ $i=\overline{1,n}$
Block (or element) identifier	(1)	$(a,1)=(x^t,1)$	$(a,b,1)$ $=(x^{j-1},x^{i-1},1)$
Block base	((1),(0, 1))	((1),(0,a,1))	((b,1),(a,0,1))

Table 6. PP2-Block and PP2-dual block (based on the block bases) building algorithms

Algorithm **Block_PP2**(*id*)	Algorithm **Dual_Block_PP2** (*id*)
Input: *id*= [*a*, *b*, *c*]. Output: list of block elements identifiers	Input: *id* = [*x*,*y*,*z*]. Output: list of identifiers of block, containing given element

x=[0,1]
B=[]
if [*a*,*b*,*c*]=[1,0,0]:
$e_1 = [1,0,0]$
$e_2 = [0,1,0]$
if [*a*,*b*,*c*]=[*a*,1,0]:
$e_1 = [1,0,0]$
$e_2 = [0, a, 1]$
if [*a*,*b*,*c*]=[*a*,*b*,1]:
$e_1 = [b, 1,0]$
$e_2 = [a, 0,1]$
B.append(e_1, e_2)
$a = e_1$
for *i* in range(1,*n*):
B.append($a + e_2$)
$a = a * x$
return *B*
	D=[]
	if [*x*,*y*,*z*]=[1,0,0]:
	D.append([1,0,0])
	for *c* in *GF*(*n*):
	D.append([*c*,1,0])
	return *D*
	if [*x*,*y*,*z*]=[*x*,1,0]:
	D.append([1,0,0])
	for *c* in *GF*(*n*):
	D.append([*c*, *x*, 1])
	return *D*
	if [*x*,*y*,*z*]=[*x*,*y*,1]:
	D.append([*y*,1,0])
	for *c* in *GF*(*n*):
	D.append([*x*-*y*∗*c*,*c*,1])
	return *D*

Table 9 gives algorithms for constructing blocks and dual blocks given by their identifiers [*a*, *b*, *c*,*d*] and [*w*,*x*,*y*,*z*]. Table 10 represents algorithms for the intersection a pair of blocks specified by identifiers [a_i, b_i, c_i, d_i], [a_j, b_j, c_j, d_j] and the intersection a pair of dual blocks specified by identifiers [w_i, x_i, y_i, z_i], [w_j, x_j, y_j, z_j].. Parameters *p*, *l*, *n* are global.

3.4 Algorithms for Projective Planes PP(2,*n*) Based on the Balancing of Element and Block Identifiers

The elements also form a projective space of dimension three and are represented by normalized polynomials of degree at most three (1), (*y*,1), (*x*,*y*,1). The general notation of element identifier is taken as (*x*,*y*,*z*) ∈ GF(p^n). Blocks also correspond to elements of the projective space and are represented by sets of coefficients of normalized polynomials written in inverse order (1,*b*,*c*), (0,1,*c*),(1). The common notation of block identifier is taken as (*a*,*b*,*c*) ∈ GF(p^n).The relation $\{((x,y,z),(a,b,c)):(ax + by + cz) = 0\}$ on Cartesian pruduct $\mathbf{X} \times \mathbf{B}$ is used, corresponding to the entry of the element with identifier (*x*,*y*,*z*) into the block with identifier (*a*,*b*,*c*)[].

Table 7. PP2-Blocks and PP2-dual blocks (based on the block bases) intersection building algorithms

Algorithm **Blocks_PP2_Intersection** (*id*1,*id*2)	Algorithm **Dual_Blocks_ PP2_Intersection**(*id*1,*id*2)
Input: *id*1=[a_i, b_i, c_i], *id*2=[a_j, b_j, c_j] Output: common element identifier	Input: *id*1 = [x_i, y, z_i], *id*2=[x_j, y_j, z_j] Output: identifier of block, containing pair of given elements

if$[a_i, b_i, c_i] = [1,0,0]$ and $[a_j, b_j, c_j] =$
$= [a_j, 1, 0]$ or $[a_i, b_i, c_i] =$
$[a_i, 1,0]$ and $[a_j, b_j, c_j] = [a_j, 1, 0]$:
return [1, 0, 0]
if $[a_i, b_i, c_i] = [1,0,0]$ and
$[a_j, b_j, c_j] = [a_j, b_j, 1]$:
return $[b_j, 1, 0]$
if $[a_i, b_i, c_i] = [a_i, 1,0]$ and
$[a_j, b_j, c_j] = [a_j, b_j, 1]$:
return $[b_j a_i + a_j, a_i, 1]$
if $[a_i, b_i, c_i] = [a_i, b_i, 1]$ and
$[a_j, b_j, c_j] = [a_j, b_j, 1]$:
if $b_i \neq b_j$:
$s = \frac{a_j - a_i}{b_i - b_j}$
return $[b_i s + a_i, s, 1]$
else:
return $[b_j, 1, 0]$
	if $[x_i, y_i, z_i] = [1,0,0]$ and
	$[x_j, y_j, z_j] = [x_j, 1]$ or
	$[x_i, y_i, z_i] = [x_i, 1,0]$ and
	$[x_j, y_j, z_j] = [x_j, 1]$:
	return [1, 0, 0]
	if $[x_i, y_i, z_i] = [1,0,0]$ and
	$[x_j, y_j, z_j] = [x_j, x_j, 1]$:
	return $[y_j,1,0]$
	if $[x_i, y_i, z_i] = [x_i, 1,0]$ and
	$[x_j, y_j, z_j] = [x_j, y_j, 1]$:
	return $[x_j - y_j x_i, x_i, 1]$
	if $[x_i, y_i, z_i] = [x_i, x_i, 1]$ and
	$[x_j, y_j, z_j] = [x_j, x_j, 1]$:
	if $y_i \neq y_j$:
	$b = \frac{x_j - x_i}{y_j - y_i}$
	$a = x_j - b y_j$
	return $[a, b, 1]$
	else:
	return $[y_i,1,0]$

Table 8. PP3 block bases

Block (or element) number	*t*=0	*t*	*t*=*jn*+*i*, $j = \overline{1,n}$ $i = \overline{1,n}$	*t*=*jn*^{2}+*in*+*r*, $j = \overline{1,n}$ $i = \overline{1,n}$
Block (or element) identifier	(1)	$(a, 1)$=(x^t,1)	$(a, b, 1)$ $= (x^{j-1}, x^{i-1}, 1)$	$(a, b, c, 1) =$ $= (x^{r-1}, x^{j-1}, x^{i-1}, 1)$
Block base	((1), (0,1), (0,0,1))	((1), (0,1), $(0,0, a, 1)$)	((1), $(0, b, 1)$, $(0, a, 0,1)$)	((c, 1), $(b, 0,1)$, $(a, 0,0,1)$)

Table 9. PG3-Block and PG3-dual block (based on the block bases) building algorithms

Algorithm **Block_PG3**(*id*)	Algorithm **Dual_Block_PG3** (*id*)
Input: *id*= [*a*, *b*, *c, d*]. Output: list of block elements identifiers	Input: *id* = [*w,x,y,z*]. Output: list of identifiers of block, containing given element

B = []
x = [0,1]
if [*a,b,c,d*]=[1,0,0,0] or
[*a,b,c,d*]=[*a*,1,0,0]:
$e_1 = [1,0,0,0]$
$e_2 = [0,1,0,0]$
if [*a,b,c,d*]=[*a,b*,1,0]:
$e_1 = [1,0,0,0]$
$e_2 = [0, b, 1,0]$
if [*a,b,c,d*]=[*a,b,c*,1]:
$e_1 = [c, 1,0,0]$
$e_2 = [b, 0,1,0]$
B.append(e_1, e_2)
$a = e_1$
for *i* in range(1,*n*):
B.append($a + e_2$)
$a = a * x$
if [*a,b,c,d*]=[1,0,0,0]:
$e_3 = [0,0,1,0]$
if [*a,b,c,d*]=[*a*,1,0,0]:
$e_3 = [0,0, a, 1]$
if [*a,b,c,d*]=[*a,b*,1,0]:
$e_3 = [0, a, 0,1]$
if [*a,b,c,d*]=[*a,b,c*,1]:
$e_3 = [a, 0,0,1]$
B.append(e_3)
for *i* in range(0,*n*+1):
a = *B*[*i*]
for *j* in range(1,*n*):
B.append($a + e_3$)
a = *a* * *x*
return *B*
	D = []
	if [*w,x,y,z*]=[1,0,0,0]:
	D.append([1,0,0,0])
	for *c* in *GF*(*n*):
	D.append([*c*,1,0,0])
	for *c* in *GF*(*n*):
	for *d* in *GF*(*n*):
	D.append([*c*,*d*,1,0])
	return *D*
	if [*w,x,y,z*]=[*w*,1,0,0]:
	D.append([1,0,0,0])
	for *c* in *GF*(*n*):
	D.append([*c*,1,0,0])
	for *c* in *GF*(*n*):
	for *d* in *GF*(*n*):
	D.append([*c*,*d*,*w*,1])
	return *D*
	if [*w,x,y,z*]=[*w,x*,1,0]:
	D.append([1,0,0,0])
	for *c* in *GF*(*n*):
	D.append([*c*,*x*,1,0])
	for *c* in *GF*(*n*):
	for *d* in *GF*(*n*):
	D.append([*c*,*w*-*xd*,*d*,1])
	return *D*
	if [*w,x,y,z*]=[*w,x,y*,1]:
	D.append([*y*,1,0,0])
	for *c* in *GF*(*n*):
	D.append([*x* - *yc*,*c*,1,0])
	for *c* in *GF*(*n*):
	$a = y * c$
	for *d* in *GF*(*n*):
	D.append([*w* – *xd* - *a*,*c*,*d*,1])
	return *D*

Table 10. PG3-Block and PG3-dual block (based on the block bases) intersection building algorithms

Algorithm **Blocks_PG3_Intersection** (*id*1,*id*2)	Algorithm **Dual_Blocks_ PG3_Intersection**(*id*1,*id*2)
Input: *id*1=[a_i, b_i, c_i], *id*2=[a_j, b_j, c_j] Output: common element identifier	Input: *id*1 = [x_i, y, z_i], *id*2=[x_j, y_j, z_j] Output: identifier of block, containing pair of given elements

rez = []
if $[a_i, b_i, c_i, d_i] = [1,0,0,0]$ *and* $[a_j, b_j, c_j, d_j] = [a_j, 1,0,0]$ *or* $[a_i, b_i, c_i, d_i] = [a_i, 1,0,0]$ *and* $[a_j, b_j, c_j, d_j] = [a_j, 1,0,0]$:
rez.append([1,0,0,0])
for *y* in *GF*(*n*):
rez.append([*y*,1,0,0])
return *rez*
	rez = []
	if $[w_i, x_i, y_i, z_i] = [1,0,0,0]$ *and* $[w_j, x_j, y_j, z_j] = [w_j, 1,0,0]$ *or* $[w_i, x_i, y_i, z_i] = [w_i, 1,0,0]$ *and* $[w_j, x_j, y_j, z_j] = [w_j, 1,0,0]$:
	rez.append([1,0,0,0])
	for *e* in *GF*(*n*):
	rez.append([*e*,1,0,0])
	return *rez*

if $[a_i, b_i, c_i, d_i] = [1,0,0,0]$ *and* $[a_j, b_j, c_j, d_j] = [a_j, b_j, 1,0]$:
rez.append([1,0,0,0])
for *y* in *GF*(*n*):
rez.append([*y*,b_j,1,0])
return *rez*
	if $[w_i, x_i, y_i, z_i] = [1,0,0,0]$ *and* $[w_j, x_j, y_j, z_j] = [w_j, x_j, 1,0]$:
	rez.append([1,0,0,0])
	for *e* in *GF*(*n*):
	rez.append([*e*,x_j,1,0])
	return *rez*

if $[a_i, b_i, c_i, d_i] = [1,0,0,0]$ *and* $[a_j, b_j, c_j, d_j] = [a_j, b_j, c_j, 1]$:
rez.append([c_j,1,0,0])
for *y* in *GF*(*n*):
rez.append([$c_j y + b_j$,*y*,1,0])
return *rez*
	if $[w_i, x_i, y_i, z_i] = [1,0,0,0]$ *and* $[w_j, x_j, y_j, z_j] = [w_j, x_j, y_j, 1]$:
	rez.append([y_j,1,0,0])
	for *e* in *GF*(*n*):
	rez.append([$x_j - y_j e$,*e*,1,0])
	return *rez*

if $[a_i, b_i, c_i, d_i] = [a_i, 1,0,0]$ *and* $[a_j, b_j, c_j, d_j] = [a_j, b_j, 1,0]$:
rez.append([1,0,0,0])
$s = a_i b_j + a_j$
for *y* in *GF*(*n*):
rez.append([*y*,*s*,a_i,1])
return *rez*
	if $[w_i, x_i, y_i, z_i] = [w_i, 1,0,0]$ *and* $[w_j, x_j, y_j, z_j] = [w_j, x_j, 1,0]$:
	rez.append([1,0,0,0])
	$b = w_j - w_i x_j$
	for *e* in *GF*(*n*):
	rez.append([*e*,*b*,w_i,1])
	return *rez*

if $[a_i, b_i, c_i, d_i] = [a_i, 1,0,0]$ *and* $[a_j, b_j, c_j, d_j] = [a_j, b_j, c_j, 1]$:
rez.append([c_j,1,0,0])
$s = a_i b_j + a_j$
for *y* in *GF*(*n*):
rez.append([$c_j y + s$,*y*,a_i,1])
return *rez*
	if $[w_i, x_i, y_i, z_i] = [w_i, 1,0,0]$ *and* $[w_j, x_j, y_j, z_j] = [w_j, x_j, y_j, 1]$:
	rez.append([y_j,1,0,0])
	$a = w_j - w_i x_j$

if	

(*continued*)

Table 10. (*continued*)

$[a_i, b_i, c_i, d_i] =$
$[a_i, b_i, 1,0]$ *and* $[a_j, b_j, c_j, d_j] =$
$[a_j, b_j, 1,0]$:
rez.append([1,0,0,0])
if $b_i \neq b_j$:
$c = \frac{a_j - a_i}{b_i - b_j}$
$b = c * b_i + a_i$
for *y* in *GF*(*n*):
if $b_i \neq b_j$:
rez.append([*y*,*b*,*c*,1])
else:
rez.append($[y, b_j, 1,0]$)
return *rez*
if $[a_i, b_i, c_i, d_i] = [a_i, b_i, 1,0]$ *and*
$[a_j, b_j, c_j, d_j] = [a_j, b_j, c_j, 1]$:
rez.append($[c_j b_i + b_j, b_i, 1,0]$)
for *y* in *GF*(*n*):
$b = b_i y + a_i$
rez.append($[c_j b + b_j y + a_j, b, y, 1]$)
return *rez*
if $[a_i, b_i, c_i, d_i] = [a_i, b_i, c_i, 1]$ *and*
$[a_j, b_j, c_j, d_j] = [a_j, b_j, c_j, 1]$:
if $b_j = b_i$ and $c_j = c_i$:
rez.append($[c_j, 1,0,0]$)
for *y* in *GF*(*n*):
rez.append($[c_j y + b_j, y, 1,0]$)
return *rez*
elif $c_j = c_i$:
rez.append($[c_j, 1,0,0]$)
$s_1 = \frac{a_i - a_j}{b_j - b_i}$
$s_2 = b_j s_1 + a_j$
for *y* in *GF*(*n*):
rez.append($[c_j y + s_2, y, s_1, 1]$)
return *rez*
else:
$inv_c = \frac{1}{c_j - c_i}$
$s_2 = (b_i - b_j) * inv_c$
rez.append($[c_j s_2 + b_j, s_2, 1,0]$)
$s_3 = (a_i - a_j) * inv_c$
for *e* in *GF*(*n*):
rez.append($[a - y_j e, e, w_i, 1]$)
return *rez*
if $[w_i, x_i, y_i, z_i] = [w_i, x_i, 1,0]$
and $[w_j, x_j, y_j, z_j] =$
$[w_j, x_j, 1,0]$:
rez.append([1,0,0,0])
if $x_i \neq x_j$:
$b = \frac{w_j - w_i}{x_j - x_i}$, $a = w_j - b x_j$
for *e* in *GF*(*n*):
if $x_i \neq x_j$:
rez.append($[e, x_j, 1,0]$)
else:
rez.append($[e, a, b, 1]$)
return *rez*
if $[w_i, x_i, y_i, z_i] = [w_i, x_i, 1,0]$
and $[w_j, x_j, y_j, z_j] =$
$[w_j, x_j, y_j, 1]$:
$s_1 = x_j - x_i y_j$
rez.append($[s_1, x_i, 1,0]$)
$s_2 = w_j + w_i y_j$
for *e* in *GF*(*n*):
rez.append($[s_2 - s_1 e, w_i - x_i e, e, 1]$)
return *rez*

if $[w_i, x_i, y_i, z_i] = [w_i, x_i, y_i, 1]$
and $[w_j, x_j, y_j, z_j] =$
$[w_j, x_j, y_j, 1]$:
if $x_j = x_i$ *and* $y_j = y_i$:
rez.append($[y_j$,1,0,0])
for *e* in *GF*(*n*):
rez.append($[x_j - y_j e, e, 1,0]$)
return *rez*
elif $y_j = y_i$:
rez.append($[y_j$,1,0,0])
$inv_x = \frac{1}{x_i - x_j}$

(*continued*)

Table 10. (*continued*)

for y in $GF(n)$:
$s_1 = s_3 + s_2 y$
rez.append($[c_j s_1 + b_j y + a_j, s_1, y, 1]$)
return rez
	$sub_w = (w_i - w_j) * inv_x$
	$sub_y = (y_j - y_i) * inv_x$
	for e in $GF(n)$:
	$c = sub_w + e * sub_y$
	$a = w_j - c * x_j - y_j e$
	rez.append($[a, e, c, 1]$)
	return rez
	else:
	$inv_y = \frac{1}{y_i - y_j}$
	$sub_x = (x_i - x_j) * inv_y$
	$a = x_j - sub_x * y_j$
	rez.append($[a, sub_x, 1,0]$)
	$sub_w = (w_i - w_j) * inv_y$
	for e in $GF(n)$:
	$b = sub_w + sub_x * e$
	rez.append($[w_j - x_j e - y_j b, b, e, 1]$)
	return rez

Algorithms **Block_PP2a**(id) and **Dual_Block_PP2a**(id) for construction a block and a dual block of the projective plane, as well as algorithms **Blocks_PP2a_Intersection** ($id1$,$id2$) and **Dual_Blocks_ PP2a_Intersection**($id1$,$id2$) for intersection of blocks and intersection of dual blocks of the projective plane by balancing block identifiers and element identifiers were conducted in [13, 17].

3.5 Algorithms for Projective Geometries PG(3,*n*) Based on the Balancing of Element and Block Identifiers

The elements also form a projective space of dimension four and are represented by normalized polynomials of degree at most four (1), (y,1), (x,y,1), (w,x,y,1). The general notation of element identifier is taken as (w,x,y,z) ∈ GF(n^4). Blocks also correspond to elements of the projective space and are represented by sets of coefficients of normalized polynomials written in inverse order (1,,b,c,d), (0,1,c,c),(1,d),(1). The common notation of block identifier is taken as (a,b,c.d) ∈ GF(n^4).

The relation $\{((w,x,y,z),(a,b,c,d)):(aw + bx + cy + dz) = 0\}$ on Cartesian product $\tilde{\boldsymbol{X}} \times \tilde{\boldsymbol{B}}$ is used, corresponding to the entry of the element with identifier (x,y,z) into the block with identifier (a,b,c)[].

Table 11 gives algorithms for constructing blocks and dual blocks given by their identifiers [a, b, c] and [x,y,z]. Table 12 represents algorithms for the intersection a pair

of blocks specified by identifiers $[a_i, b_i, c_i]$, $[a_j, b_j, c_j]$ and the intersection a pair of dual blocks specified by identifiers $[x_i, y_i, z_i]$, $[x_j, y_j, z_j]$. Parameters *p, l, n* are global.

Table 11. PG3-Block and PG3-dual block (based on the balancing of element and block identifiers) building algorithms

Algorithm **Block_PP3a**(*id*)	Algorithm **Dual_Block_PP3a**(*id*)
Input: *id*= [*a*, *b*, *c, d*]. Output: list of block elements identifiers	Input: *id* = [*w,x,y,z*]. Output: list of identifiers of block, containing given element

B = []
if [*a*, *b*, *c*, *d*] = [0, 0, 0, 1]:
for *w* in *GF*(*n*):
B.append([*w*, 1, 0, 0])
for *x* in *GF*(*n*):
B.append([*w*, *x*, 1, 0])
B.append([1, 0, 0, 0])
if [*a*, *b*, *c*, *d*] = [0, 0, 1, *d*]:
for *w* in *GF*(*n*):
B.append([*w*, 1, 0, 0])
for *x* in *GF*(*n*):
B.append([*w*, *x*, –*d*, 1])
B.append([1, 0, 0, 0])
if [*a*, *b*, *c*, *d*] = [0, 1, *c*, *d*]:
for *w* in *GF*(*n*):
B.append([*w*, –*c*, 1, 0])
for *y* in *GF*(*n*):
B.append([*w*, –*cy*–*d*, *y*, 1])
B.append([1, 0, 0, 0])
if [*a*, *b*, *c*, *d*] = [1, *b*, *c*, *d*]:
for *x* in *GF*(*n*):
B.append([–*bx*–*c*, *x*, 1, 0])
for *y* in *GF*(*n*):
B.append([–*bx*–*cy*–*d*, *x*, *y*, 1])
B.append([–*b*, 1, 0, 0])
return *B*
	D = []
	if [*w*, *x*, *y*, *z*] = [1, 0, 0, 0]:
	for *d* in *GF*(*n*):
	D.append([0, 0, 1, *d*])
	for *c* in *GF*(*n*):
	D.append([0, 1, *c*, *d*])
	D.append([0, 0, 0, 1])
	if [*w*, *x*, *y*, *z*] = [*w*, 1, 0, 0]:
	for *d* in *GF*(*n*):
	D.append([0, 0, 1, *d*])
	for *c* in *GF*(*n*):
	D.append([1, –*w*, *c*, *d*])
	D.append([0, 0, 0, 1])
	if [*w*, *x*, *y*, *z*] = [*w*, *x*, 1, 0]:
	for *d* in *GF*(*n*):
	D.append([0, 1, –*x*, *d*])
	for *b* in *GF*(*n*):
	D.append([1,*b*,–*xb*–*w*, *d*])
	D.append([0, 0, 0, 1])
	if [*w*, *x*, *y*, *z*] = [*w*, *x*, *y*, 1]:
	for *c* in *GF*(*n*):
	D.append([0, 1, *c*, –*yc*–*x*])
	for *b* in *GF*(*n*):
	D.append([1,*b*,*c*,–*yc*–*xb*–*w*])
	D.append([0, 0, 1, –*y*])
	return *D*

Table 12. PG3-blocks and PG#-dual blocks (based on the balancing of element and block identifiers) intersection building algorithms

Algorithm **Blocks_PP3a_Intersection** (*id*1,*id*2)	Algorithm **Dual_Blocks_ PP3a_Intersection**(*id*1,*id*2)
Input: *id*1=[a_1, b_1, c_1, d_1], *id*2=[a_2, b_2, c_2, d_2] Output: common element identifier	Input: *id*1=[w_1,x_1, y_1, z_1], *id*2=[w_2,x_2, y_2, z_2] Output: identifier of block, containing pair of given elements

rez = []
if $a_i = a_j = 0$ and $b_i = b_j = 0$:
for *w* in *GF*(*n*):
rez.append([*w*, 1, 0, 0])
rez.append([1, 0, 0, 0])
if $a_i = a_j = 0$ and ($b_i = 1$ and $b_j = 0$ and $c_j = 0$ or $b_i = b_j = 1$ and $c_i = c_j$):
for *w* in *GF(n)*:
rez.append([*w*, $-c_i$, 1, 0])
rez.append([1, 0, 0, 0])
if $a_i = a_j = 0$ and $b_j = 1$ and $b_i = 0$ and $c_i = 0$:
for *w* in *GF(n)*:
rez.append([*w*, $-c_j$, 1, 0])
rez.append([1, 0, 0, 0])
if $a_i = a_j = 0$ and $b_i = 1$ and $b_j = 0$ and $c_j = 1$:
for *w* in *GF(n)*:
rez.append([*w*, $c_i d_j - d_i$, $-d_j$, 1])
rez.append([1, 0, 0, 0])
if $a_i = a_j = 0$ and $b_j = 1$ and $b_i = 0$ and $c_i = 1$:
for *w* in *GF(n)*:
rez.append([*w*, $c_j d_i - d_j$, $-d_i$, 1])
rez.append([1, 0, 0, 0])
if $a_i = a_j = 0$ and $b_j = b_i = 1$ and $c_i \neq c_j$:
$y = \frac{d_i - d_j}{c_j - c_i}$
for *w* in *GF(n)*:
rez.append([*w*, $-c_j y - d_j$, *y*, 1])
rez.append([1, 0, 0, 0])
if $a_i = 1$ and $a_j = b_j = c_j = 0$:
	rez = []
	if $z_i = z_j = 0$ and $y_i = y_j = 0$:
	for *d* in *GF*(*n*):
	rez.append([0, 0, 1, *d*])
	rez.append([0, 0, 0, 1])
	if $z_i = z_j = 0$ and [$y_i = 1$ and $y_j = 0$ and $x_j = 0$ or $y_i = y_j = 1$ and $x_i = x_j$):
	for *d* in *GF*(*n*):
	rez.append([0, 1, $-x_i$, *d*])
	rez.append([0, 0, 0, 1])
	if $z_i = z_j = 0$ and $y_j = 1$ and $y_i = 0$ and $x_i = 0$:
	for *d* in *GF*(*n*):
	rez.append([0, 1, $-x_j$, *d*])
	rez.append([0, 0, 0, 1])
	if $z_i = z_j = 0$ and $y_i = 1$ and $y_j = 0$ and $x_j = 1$:
	for *d* in *GF*(*n*):
	rez.append([1, $-w_j$, $x_i w_j - w_i$, *d*])
	rez.append([0, 0, 0, 1])
	if $z_i = z_j = 0$ and $y_j = 1$ and $y_i = 0$ and $x_i = 1$:
	for *d* in *GF(n)*:
	rez.append([1, $-w_i$, $x_j w_i - w_j$, *d*])
	rez.append([0, 0, 0, 1])
	if $z_i = z_j = 0$ and $y_j = y_i = 1$ and $x_i \neq x_j$:
	$b = \frac{w_i - w_j}{x_j - x_i}$
	for *d* in *GF*(*n*):
	rez.append([1, *b*, $-x_j b - w_j$, *d*])
	rez.append([0, 0, 0, 1])
	if $z_i = 1$ and $z_j = y_j = x_j = 0$:

(*continued*)

Table 12. (*continued*)

for x in *GF(n)*:
rez.append([$-b_i x - c_i$, x, 1, 0])
rez.append([$-b_i$, 1, 0, 0])
if $a_i = 1$ and $a_j = b_j = 0$ and $c_j = 1$:
$y = -d_j$
for x in *GF(n)*:
rez.append([$-b_i x - c_i y - d_i$, x, y, 1])
rez.append([$-b_i$, 1, 0, 0])
if $a_i = 1$ and $a_j = 0$ and $b_j = 1$:
for y in *GF(n)*:
$x = -c_j y - d_j$
rez.append([$b_i x - c_i y - d_i$, x, y, 1])
rez.append([$b_i c_j - c_i$, $-c_j$, 1, 0])
if $a_j = 1$ and $a_i = b_i = c_i = 0$:
for x in *GF(n)*:
rez.append([$-b_j x - c_j$, x, 1, 0])
rez.append([$-b_j$, 1, 0, 0])
if $a_j = 1$ and $a_i = b_i = 0$ and
rez.append([$-b_j x - c_j y - d_j$, x, y, 1])
rez.append([$-b_j$, 1, 0, 0])
if $a_j = 1$ and $a_i = 0$ and $b_i = 1$:
for y in *GF(n)*:
$x = -c_i y - d_i$
rez.append([$-b_j x - c_j y - d_j$, x, y, 1])
rez.append([$b_j c_i - c_j$, $-c_i$, 1, 0])
if $a_i = a_j = 1$ and $b_i = b_j$:
rez.append([$-b_i$, 1, 0, 0])
if $a_i = a_j = 1$ and $b_i = b_j$ and $c_i = c_j$:
for x in *GF(n)*:
rez.append([$-b_i x - c_i$, x, 1, 0])
if $a_i = a_j = 1$ and $b_i \neq b_j$:
$x = \frac{c_i - c_j}{b_j - b_i}$
rez.append([$-b_i x - c_i$, x, 1, 0])
for c in *GF(n)*:
rez.append([0,1,c,$-y_i c - x_i$])
rez.append([0, 0, 1, $-y_i$])
if $z_i = 1$ and $z_j = y_j = 0$ and $x_j = 1$:
$b = -w_j$
for c in *GF(n)*:
rez.append([1, b, c, $-y_i c - x_i b - w_i$])
rez.append([0, 0, 1, $-y_i$])
if $z_i = 1$ and $z_j = 0$ and $y_j = 1$:
for b in *GF(n)*:
$c = -x_j b - w_j$
rez.append([1, b, c, $y_i c - x_i b - w_i$])
for d in *GF(n)*:
rez.append([0,1,$-x_j$,$y_i x_j - x_i$])
if $z_j = 1$ and $z_i = y_i = x_i = 0$:
for c in *GF(n)*:
rez.append([0,1,c, $-y_j c - x_j$])
rez.append([0, 0, 1,$-y_j$])
if $z_j = 1$ and $z_i = y_i = 0$ and $x_i = 1$:
$b = -w_i$
for c in *GF(n)*:
rez.append([1, b, c, $-y_j c - x_j b - w_j$])
rez.append([0, 0, 1, $-y_j$])
if $z_j = 1$ and $z_i = 0$ and $y_i = 1$:
for b in *GF(n)*:
$c = -x_i b - w_i$
rez.append([1, b, c, $-y_j c - x_j b - w_j$])
rez.append([0, 1, $-x_i$, $y_j x_i - x_j$])
if $z_i = z_j = 1$ and $y_i = y_j$:
rez.append([0, 0, 1, $-y_i$])
if $z_i = z_j = 1$ and $y_i = y_j$ and $x_i = x_j$:
for c in *GF(n)*:

(*continued*)

Table 12. (*continued*)

if $a_i = a_j = 1$ and $b_i = b_j$ and $c_i \neq c_j$:
  $y = \frac{d_i - d_j}{c_j - c_i}$
  for x in GF(n):
    rez.append([$-b_i x - c_i y - d_i$, x, y, 1])
if $a_i = a_j = 1$ and $b_i \neq b_j$:
  $inv = \frac{1}{b_j - b_i}$
  $[m_1, m_2] = [(c_i - c_j) * inv, (d_i - d_j) * inv]$
  for y in GF(n):
    $x = m_1 y + m_2$
    rez.append([$-b_i x - c_i y - d_i$, x, y, 1])
return rez

    rez.append([0, 1, x, $-y_i c - x_i$])
if $z_i = z_j = 1$ and $y_i \neq y_j$:
  $c = \frac{x_i - x_j}{y_j - y_i}$
  rez.append([0, 1, c, $-y_i c - x_i$])
if $z_i = z_j = 1$ and $y_i = y_j$ and $x_i \neq x_j$:
  $b = \frac{w_i - w_j}{x_j - x_i}$
  for c in *GF*(n):
    rez.append([1, b, c, $-y_i c - x_i b - w_i$])
if $z_i = z_j = 1$ and $y_i \neq y_j$:
  $inv = \frac{1}{y_j - y_i}$
  $[m_1, m_2] = [(x_i - x) inv, (w_i - w_j) inv]$
  for b in *GF*(n):
    $c = m_1 b + m_2$
    rez.append([1, b, c, $-y_i c - x_i b - w_i$])
return *rez*

3.6 Algorithms for Cyclic Projective Planes PP(2,*n*) and

(v,k,1)-Cyclic projective plane PP(2,n) is defined by (v,k,1)-difference Zinger set $d = (0,d_1,\ldots, d_{n-1})$. ($v$,$k$,3)-Cyclic projective geometry PG(3,n) is defined by (v,k,3)-difference Zinger set $d = (0, d_1, ..., d_{n^2+n})$ The block identifiers are their numbers $s \in \{0,\ldots,v-1\}$. The block with number s is obtained by adding its number to the elements of the difference set in the ring Z_v; the dual block is obtained by subtracting the elements of the difference set from its number in the ring Z_v. [18].

Table. 13 represents algorithms for the intersection a pair of blocks specified by identifiers s_1 and s_2 and the intersection a pair of dual blocks specified by element identifiers x_i and y_i. Parameters v, d are global.

3.7 Algorithms for Linear Transversal Combinatorial Block-Designs TD(2,*k*,*n*)

The linear transversal combinatorial block design TD(2,k,n) [3, 4] is constructed from elements of k pairwise disjoint sets w_i containing n elements in each, and consists of n^2

Table 13. Cyclic PP2-blocks, PP2-dual blocks PG3-blocks, PG3-dual blocks intercection building algorithms

Algorithm **Blocks_PP2b_Intersection** (*id*1,*id*2) Algorithm **Blocks_PP3b_Intersection** (*id*1,*id*2)	Algorithm **Dual_Blocks_ PP2b_Intersection**(*id*1,*id*2) Algorithm **Dual_Blocks_ PP3b_Intersection**(*id*1,*id*2)
Input: $id1=s_1$, $id2=s_2$ Output: common element identifier	Input: $id1=x_1$, $id2=x_2$ Output: identifier of block, containing pair of given elements

rez = []
$s = s_2 - s_1$
for *el* in *d*:
  if ((*s* + *el*) mod *v*) in *d*:
  rez.append((*el* + s_2) mod *v*)
return *rez*
	rez = []
	$x = x_2 - x_1$
	for *el* in *d*:
	  if ((*x* + *el*) mod *v*) in *d*:
	  rez.append((−*el* + x_1) mod *v*)
	return *rez*

sets $\mathbf{Y}_j$, $j=\overline{(1, n^2)}$ containing k elements in each. In this case, each set $\mathbf{Y}_j$ has exactly one element from each set $\mathbf{W}_i$, $i = 0,\ldots,k-1$, and if $j \neq s$, then $\mathbf{Y}_j$ and $\mathbf{Y}_s$ have at most one element in common. Thus, the combinatorial block design TD(k,n) is defined on the set $\mathbf{X}$ of power kn. If n is a prime or a prime power, then TD(k,n) exists.

Then in the algebraic interpretation, the set $\mathbf{X}$ is the set of pairs (y,x), $x \in \mathbf{X}_1$, where $\mathbf{X}_1$ is the set of k elements of the field F_n, $y \in F_n$; blocks of linear transversal combinatorial block design TD(k,n) are defined by pairs of elements $(a,b) \in$ GF(n^2) [3]:

$\mathbf{B}_{a,b} = \{(ax + b, x)\colon x \in \mathbf{X}_1\}$.

Let x be primitive element of $GF(n)$, $n = p^l$ where p is prime. Denote $[x,y]$ an element identifier. It is dual block identifier simultaneously. Identifiers of blocks will be denoted, $[a,b]$ $x,y,a,b \in GF(n)$ [LS].

Table 14 gives algorithms for constructing blocks and dual blocks defined by their identifiers [*a, b*] and [*x, y*]. Table 15 represents algorithms for operations of the intersection a pair of blocks specified by identifiers [1], [2] and the intersection a pair of dual blocks specified by identifiers $[x_1, y_1]$ and $[x_2,y_2]$. Parameters *k, n* are global.

Table 14. TD2-Block and TD2-dual block building algorithms

Algorithm **Block_TD2**(*id*)	Algorithm **DualBlock_TD2** (*id*)
Input: *id*= [*a*, *b*]. Output: list of block elements identifiers	Input: *id* = [*x*,*y*]. Output: list of identifiers of block, containing given element

B=[[*b*,0]]
$x_0 = x$
for *t* in range(0,*k*):
B.append($[ax_0 + b, x_0]$)
$x_0 = x_0 * x$
return *B*
	D=[]
	for *a* in GF(*n*):
	if *y* = 0:
	b = *x*
	else:
	b=*x*–*ay*
	D.append([*a*,*b*])
	return *D*

3.8 Algorithms for Quadratic Transversal Combinatorial Block-Designs TD(3,*k*,*n*)

The quadric transversal combinatorial block design TD(3,*k*,*n*) [3, 4] is constructed from elements of *k* pairwise disjoint sets w_i containing *n* elements in each, and consists of n^3 sets $\mathbf{Y}_j$, $j=\overline{(1, n^2)}$ containing *k* elements each. In this case, each set $\mathbf{Y}_j$ has exactly one element from each set $\mathbf{W}_i$, $i = 0,\ldots,k-1$, and if $j \neq s$, then $\mathbf{Y}_j$ and $\mathbf{Y}_s$ have at most two element in common. Thus, the combinatorial block design TD(3,*k*,*n*) is defined on the set **X** of power *kn*. If *n* is a prime or a prime power, then TD(3,*k*,*n*) exists.

Table 15. TD2-Block and TD2-dual block intercection building algorithms

Algorithm **Blocks_TD2_Intersection** (*id*1,*id*2)	Algorithm **Dual_Blocks_ TD2_Intersection**(*id*1,*id*2)
Input: *id*1=[a_1, b_1], *id*2=[a_2, b_2] Output: common element identifier	Input: *id*1=[x_1, y_1], *id*2=[x_2, y_2] Output: identifier of block, containing pair of given elements

if $a_1 \neq a_2$:
if $b_1 = b_2$:
x=0
else:
$x = \frac{b_2 - b_1}{a_1 - a_2}$
if *x*=0 or ind(*x*) < *k*:
return $[a_1x + b_1, x]$
return []
	if $y_1 \neq y_2$:
	if $x_1 = x_2$:
	a=0
	b=x_1
	else:
	$a = \frac{x_1 - x_2}{y_1 - y_2}$
	$b = x_1 - ay_1$
	return [*a*,*b*]
	return []

Then in the algebraic interpretation, the set $\mathbf{X}$ is the set of pairs (y,x), $x \in \mathbf{X}_1$, where $\mathbf{X}_1$ is the set of k elements of the field F_n, $y \in F_n$; blocks of linear transversal combinatorial block design TD(3,k,n) are defined by triple of elements $(a,b,c) \in \mathrm{GF}(n^3)$ [3]:

$\mathbf{B}_{a,b} = \{(ax + b, x)\colon x \in \mathbf{X}_1\}$.

Let x be primitive element of $GF(n)$, $n = p^l$ where p is prime. Denote $[x,y]$ an element identifier. It is dual block identifier simultaneously. Identifiers of blocks will be denoted $[a,b,c]$. $x,y,a,b,c \in GF(n)$ [LS].

Table 16 gives algorithms for constructing blocks and dual blocks defined by their identifiers $[a, b]$ and $[x, y]$. Table 17 represents algorithms for operations of the intersection a pair of blocks specified by identifiers $[a_1, b_1, c_1]$,], $[a_2, b_2, c_2]$ and the intersection a pair of dual blocks specified by identifiers $[x_1, y_1]$ and $[x_2,y_2]$. Parameters k, n are global.

Table 16. TD3-Block and TD3-dual block building algorithms

Algorithm **Block_TD3**(*id*)	Algorithm **DualBlock_TD3** (*id*)
Input: *id*= [*a*, *b*, *c*]. Output: list of block elements identifiers	Input: *id* = [*x*,*y*]. Output: list of identifiers of block, containing given element

B=[]
$x_0 = x$
for *i* in range(0,*k*):
  B.append($[ax_0^2 + bx_0 + c, x_0]$)
  $x_0 = x_0 * x$
return *B*
	D=[]
	$y_{sqr} = y^2$
	for *a* in *GF*(*n*):
	  for *b* in *GF*(*n*):
	  if *y* = 0:
	  c = *x*
	  else:
	  $c = x - ay_{sqr} - by$
	  D.append([*a*,*b*,*c*])
	return *D*

4 Discussion: About Using Basic Algorithms

4.1 Algorithms for Affine Geometries

Consider affine planes RPG(2,*n*) (Reduced Projective Geometries(2,*n*)). Let us accept the identifier of an affine plane of order *n* as the identifier I_B of reduced block of projective geometry of order *n*. Later let us accept the block identifier of the affine plane of order *n* as the pair $(I_{B'}, I_B)$, $I_{B'} \neq I_B$.

Table 17. TD3-Blocks and TD3-dual blocks intercection building algorithms

Algorithm **Blocks_TD3_Intersection** (*id*1,*id*2)	Algorithm **Dual_Blocks_ TD3_Intersection**(*id*1,*id*2)
Input: *id*1=[a_1, b_1, c_1], *id*2=[a_2, b_2, c_2] Output: common element identifier	Input: *id*1=[x_1, y_1], *id*2=[x_2, y_2] Output: identifier of block, containing pair of given elements

rez = []
if $a_1 = a_2, b_1 \neq b_2, c_1 = c_2$:
rez.append($[c_1, 0]$)
if $a_1 = a_2, b_1 \neq b_2, c_1 \neq c_2$:
$x = \frac{c_2 - c_1}{b_1 - b_2}$
rez.append($[a_1x^2 + b_1x + c, x]$)
if $a_1 \neq a_2, b_1 = b_2, c_1 = c_2$:
rez.append($[c_1, 0]$)
if $a_1 \neq a_2, b_1 = b_2, c_1 \neq c_2$:
$x_{1,2} = \pm\sqrt{\frac{c_2 - c_1}{a_1 - a_2}}$
rez.append($[a_1x_1^2 + b_1x_1 + c, x_1], [a_1x_2^2 + b_1x_2 + c, x_2]$)
if $a_1 \neq a_2, b_1 \neq b_2, c_1 = c_2$:
rez.append($[c_1, 0]$)
$x = \frac{b_2 - b_1}{a_1 - a_2}$
rez.append($[a_1x^2 + b_1x + c, x]$)
if $a_1 \neq a_2, b_1 \neq b_2, c_1 \neq c_2$:
$x_{1,2} = \frac{(b_2 - b_1) \pm \sqrt{(b_1 - b_2)^2 - 4(a_1 - a_2)(c_1 - c_2)}}{2(a_1 - a_2)}$
rez.append($[a_1x_1^2 + b_1x_1 + c, x_1], [a_1x_2^2 + b_1x_2 + c, x_2]$)
return *rez*
	rez = []
	if $y_1 = y_2$:
	if $x_1 = x_2$:
	$b_1 = 0$
	else:
	$b_1 = \frac{x_1 - x_2}{y_1 - y_2}$
	$b_2 = y_1 + y_2$
	$y_{sqr} = y_1^2$
	for *a* in *GF*(*n*):
	$b = b_1 - b_2a$
	$c = x_1 - ay_{sqr} - by_1$
	rez.append([*a*,*b*,*c*])
	return *rez*

The block with identifier ($I_{B'}$,I_B) of the affine plane given by the block identifier I_B is $B'\backslash(B' \cap B)$. The dual block $D(e)$ corresponding to the element e, $e \notin B' \cap B$ of the affine plane given by the block identifier I_B is $\{(I_{B'},I_B):B' \in D(e)\}$, where $D(e)$ is the dual projective plane block.

Properties of affine plane blocks:

The intersection of the blocks given by the identifiers ($I_{B'}$,I_B) and ($I_{B''}$,I_B) of the affine plane given by the block identifier I_B of the projective geometry is ($B' \cap B''$ if B' $B' \cap B''$ $B'' \neq (B'$ $B' \cap B''$ $B) \cup (B''$ $B' \cap B''$ $B)$, and it is equal to $\emptyset$ otherwise.

The intersection of two dual blocks of the affine plane is equal to the intersection of the dual blocks of projective geometry defined by the same identifiers, because in this case, the identifiers of the elements of the block defined by the identifier of the affine plane are not used.

The dual block of the affine plane coincides with the dual block of the projective plane with the same identifier, because this identifier does not belong to the block defined by the affine plane identifier.

The algorithm for constructing a block of an affine plane differs from the algorithm for constructing a block of a projective plane in that the result of the intersection of blocks defined by the identifier of this block is removed from the result.

Using these properties and the basic algorithms discussed in the previous section, it is easy to construct algorithms **BlockRPG2**(*id*,*id*), **DualBlockRPG2**(*id,id*), **BlockRPG2Intersection**(*id*,*id*,*id*)), **DualBlockRPG2Intersection**(*id*,*id*,*id*), the detailed description of which will be omitted.

4.2 Other Compositions of Basic Algorithms

Using the above basic algorithms. More complex algorithms can be built. In routing algorithms, it is necessary to find blocks with each of which each of two non-intersecting blocks intersects. Table 18 shows the algorithm for constructing a list of such blocks in relation to STS_Bose. Algorithms for CBD STS_Skolem, TD2, TD3, RPG2, are similar, obtained from it by replacing CBD names.

Table 18. Algorithm for deriving the list of blocks intersecting with each of two non-intersecting blocks

Algorithm **Adjasent_Block_STS_Bose(***id*1,*id*2**)**
Input: $id1=[x_1, y_1, i_1]$, $id1= [x_2, y_2, i_2]$, $B_{id1} \cap B_{id1}=\varnothing$ Output: list of blocks identifiers *id* such that $B_{id} \cap B_{id1}$ $B_{id1} \cap B_{id1} \neq \varnothing$, $B_{id} \cap B_{id2} \neq \varnothing$

rez = []
B_1 = **Block_STS_Bose**($id1$)
B_2 = **BlockSTSBose**($id2$)
for $[z_1, i_1]$ in B_1:
for $[z_2, i_2]$ in B_2:
for *it* in **Dual_Block_STS_Bose_Intersection**($[z_1 , i_1]$, $[z_2 , i_2]$
if *it* not in *rez*: *rez*.append(*it*)
return *rez*

4.3 About Numerical Notations of Algorithms

All the above algorithms are given in algebraic notation with representation of inputs and outputs by identifiers. They can be used in composition with the numbering algorithms $\widetilde{\mathbf{X}} \to \mathrm{N} \cup \{0\}$ of elements and $\widetilde{\mathbf{B}} \to \mathrm{N} \cup \{0\}$ blocks, as well as algorithms $\mathrm{Im}(\widetilde{\mathbf{X}}) \to \widetilde{\mathbf{X}}$ of element numbering inversion and $\widetilde{\mathbf{B}} \to \mathrm{N} \cup \{0\}$ inverting the numbering of blocks, so that the representations of CBD and their blocks become convenient for analysis and subsequent use. In Tables 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32 and 33 algorithms for numbering and inversion the numbering of elements and blocks of combinatorial block designs are given.

Table 19. Algorithms for Bose STS-blocks numbering and inverse numbering

Algorithm **Blocks_STS_ Bose_ Numbering_Inversion**(g)	Algorithm **Blocks_STS_Bose_ Numbering**(id)
Input: block number s Output: block identifier id= [x,i]	Input: block identifier id = [a,b] Output: block number s
if s < n:return [s, s, 0] else: $i = \left\lfloor \frac{s-n}{tn} \right\rfloor$ $y = \left\lfloor \frac{1+\sqrt{(1+8\,(s-(t\,i+1)n))}}{2} \right\rfloor$ $x = (s - (ti + 1)n - \frac{y(y-1)}{2})$ return [x, y, i]	if $x = y$: return x else: $s = x + (t\,i + 1)n + \left\lfloor \frac{y(y-1)}{2} \right\rfloor$ return s

Table 20. Algorithms for Bose STS-elements numbering and inverse numbering

Algorithm **Elements_STS_ Bose_Numbering_Inversion**(g)	Algorithm **Elements_STS_ Bose_Numbering**(id)
Input: element number g Output: element identifier id= [x,i]	Input: element identifier id = [x,i] Output: element number g
$i = \left\lfloor \frac{g}{n} \right\rfloor$; x= g-i*n return [x i]	g= $ni + x$ return g

Table 21. Algorithms for Skolem STS-blocks numbering and inverse numbering

Algorithm **Blocks_STS_Skolem_ Numbering Inversion**(g)	Algorithm **Blocks_STS_Skolem_ Numbering**(id)
Input: block number s Output: block identifier id= [x,i]	Input: block identifier id = [a,b] Output: block number s

if $s < t$: return [s, s, 0]
else:
if $s < 3\left\lfloor\frac{n(n-1)}{2}\right\rfloor + t$:
$i = \left\lfloor\frac{2(s-t)}{n(n-1)}\right\rfloor$
$y = \left\lfloor\frac{1+\sqrt{\left(1+8\left(s-((n-1)i+1)t\right)\right)}}{2}\right\rfloor$
$x = s - \left((n-1)i + 1\right)t - \frac{y(y-1)}{2}$
return [x, y, i]
else:
$s_0 = 3\frac{n(n-1)}{2} + t$
$y = \left\lfloor\frac{s-s_0}{3}\right\rfloor$
$i = (s - s_0)$ % 3
x = '*infty*'
return [x, y, i]
	if $x = y$: return x
	else:
	if x != '*infty*':
	return $x + (t\ i + 1)n + \frac{y(y-1)}{2} - (i+1)t$
	else:
	return $3\frac{n(n-1)}{2} + 3y + i + t$

Table 22. Algorithms for Skolem STS-elements numbering and inverse numbering

Algorithm **Elements_STS_Skolem_ Numbering Inversion**(g)	Algorithm **Elements_STS_ Skolem_Numbering** (id)
Input: element number g Output: element identifier id= [x,i]	Input: element identifier id = [x,i] Output: element number g

if $g < 3n$:
$i = g$ // n
return [g - in, i]
else: return '*infty*'
	if id != '*infty*':
	[x, i] = id
	$g = ni + x$
	else:
	$g = 3n$
	return g

Table 23. Algorithms for PP2-blocks (based on the block bases) numbering and inverse numbering

Algorithm **Blocks_PP2_Numbering Inversion**(g)	Algorithm **Blocks_ PP2_ Numbering**(id)
Input: block number s Output: block identifier id= [x,i]	Input: block identifier id = [a,b] Output: block number s
if s = 0: return [0, 0] if s - 1 < n: return [x^{s-1}, 0] else: $m0$ = s mod n $m1$ = (s / p) mod n return [x^{m0}, x^{m1}]	match [a,a]: case [0,0]: return 0 case [a,0]: return ind(a) case [0,b]: return n*ind(b) case _: return ind(a)+n*ind(b)

Table 24. Algorithms for PP2-block (based on the block bases) elements numbering and inverse numbering

Algorithm **Elements_PP2_ Numbering Inversion**(g)	Algorithm **Elements_ PP2_ Numbering**(id)
Input: element number g Output: element identifier id= [x,i]	Input: element identifier id = [x,i] Output: element number g

if g = 0:
return [0, 0]
if g - 1 < p:
return [x^{g-1}, 0]
else:
$m0$ = g mod n
$m1$ = (g / n) mod n
return [x^{m0}, x^{m1}]
	match [x,y]:
	case [0,0]: return 0
	case [x,0]: return ind(x)
	case [0,y]: return n*ind(y)
	case_: return ind(x)+n*ind(y)

Using the algorithms for numbering and inverting the numbering of elements and blocks and algorithms in algebraic notation, one can get the numerical notations of the basic algorithms as the compositions mentioned above. The algorithms of such compositions will be denoted by the same names with the addition of the ending N and the replacement of the formal parameters at the input and output by their numbers, for example, the BlockTD2(id) and DualBlockTD2(id) algorithms correspond to the algorithms in Table 34.

We will assume that the number of an affine geometry is the number of its identifier; in the numerical interpretation, the elements of an affine geometry are given by a pair of numbers of a block identifier and an affine plane identifier. Algorithms for numbering elements of affine geometry and reversing the numbering of elements are the same

Table 25. Algorithms for PG3-blocks (based on the block bases) numbering and inverse numbering

Algorithm **Blocks_PG3_Numbering Inversion**(g)	Algorithm **Blocks_ PG3_ Numbering**(id)
Input: block number s Output: block identifier id= [x,i]	Input: block identifier id = [a,b] Output: block number s

if s = 0:
return [1, 0, 0, 0]
if s - 1 < n:
return [x^{s-1}, 1, 0, 0]
if s - 1 < n^2 + n:
m = s - n - 1
$m0$ = m mod n
$m1$ = m / n
return [x^{m0}, x^{m1}, 1]
else:
m = s - n^2 - n – 1
$m0$ = m mod n
$m1$ = (m - $m0$) mod n^2
$m2$ = m / n^2
return [x^{m0}, x^{m1}, x^{m2}, 1]
	match [a,b,c,d]:
	case [1,0,0,0]: return 0
	case [a,1,0,0]: return 1+ind(a)
	case [0,0,1,0]: return 1+n
	case [a,0,1,0]: return 1+n+ind(a)
	case [0,b,1,0]: return 1+n+ n*ind(b)
	case [a,b,1,0]: return 1+n+ind(a)+n*ind(b)
	case [0,0,0,1]: return 1+n+ n^2
	case [a,0,0,1]: return 1+n+n^2+ind(a)
	case [0,b,0,1]: return 1+n+n^2 +n*ind(b)
	case [0,0,c,1]: return 1+n+ n^2 + n^2*ind(c)
	case [a,b,0,1]:
	return 1+n+n^2 +ind(a)+n*ind(b)
	case [0,b,c,1]:
	return 1+n+n^2 +n*ind(b) + n^2*ind(c)
	case [a,0,c,1]: return
	return 1+n+n^2 +ind(a) + n^2*ind(c)
	case _:
	return 1+n+n^2+ind(a)+n*ind(b)+n^2*ind(c)

algorithms as for projective geometry (simply the numbers of element identifiers from the block defined by the identifier of the affine plane are not used).

4.4 On Algorithms Combined CBD Blocks and Dual Block Computing

Using the algorithms presented above, it is possible to compose algorithms for constructing blocks and dual blocks of combined CBD of the considered types. These algorithms will be denoted by the name CCBD, for example, let CPP2() be the algorithm for constructing all blocks of the combined projective plane, CDPP() be the algorithm for constructing all dual blocks of the combined projective plane. They are presented in Table 35.

Table 26. Algorithms for PG3-block (based on the block bases) elements numbering and inverse numbering

Algorithm **Elements_PG3_ Numbering Inversion**(g)	Algorithm **Elements_ PG3_ Numbering**(id)
Input: element number g Output: element identifier id= [x,i]	Input: element identifier id = [x,y] Output: element number g

if g = 0:
return [1, 0, 0, 0]
if g - 1 < n:
return [x^{g-1}, 1, 0, 0]
if g - 1 < n^2 + n:
m = g - n - 1
$m0$ = m mod n
$m1$ = m / n
return [$x^{m0}, x^{m1}, 1$]
else:
m = g - n^2 - n - 1
$m0$ = m mod n
$m1$ = (m - $m0$) mod n^2
$m2$ = m / n^2
return [$x^{m0}, x^{m1}, x^{m2}, 1$]
	match [w,x,y,z]:
	case [1,0,0,0]: return 0
	case [w,1,0,0]: return 1+ind(w)
	case [0,0,1,0]: return 1+n
	case [w,0,1,0]: return 1+n+ind(w)
	case [0,x,1,0]: return 1+n+ n*ind(x)
	case [w,x,1,0]: return 1+n+ind(w)+
	+n*ind(x)
	case [0,0,0,1]: return 1+n+ n^2
	case [w,0,0,1]: return 1+n+n^2+ind(w)
	case [0,x,0,1]:
	return 1+n+n^2 +n*ind(x)
	case [0,0,y,1]: return 1+n+n^2+
	+n^2*ind(y)
	case [w,x,0,1]:
	return 1+n+n^2 +ind(w)+n*ind(x)
	case [0,x,y,1]:
	return 1+n+n^2+n*ind(x)+
	n^2*ind(y)
	case [w,0,y,1]: return
	return 1+n+n^2 +ind(w) + n^2*ind(y)
	case _:
	return 1+n+n^2+ind(w)+n*ind(x)
	+n^2*ind(y)

5 Analisys: Complexity Estimation, Examples

5.1 Algorithms Complexity Estimation

The complexity of algorithms that do not use inversion is estimated as the largest number of additive and multiplicative operations in a certain algebraic structure that are spent to compute the element's identifier at the output of the algorithm or one identifier from the list of element identifiers at the output of the algorithm. The asymptotic complexity estimate for such algorithms coincides with the asymptotic complexity of the multiplication operation, because performs a certain number of multiplications. The complexity of the remaining algorithms is determined by the complexity of inversion performed once. Complexity of Algorithm **Blocks_PP2b_Intersection** ($id1$,$id2$) for cyclic projective plane (see Table 13) is k times the complexity of the addition in the ring Z_v.

Table 27. Algorithms for PP2-blocks (based on the identifiers balancing) numbering and inverse numbering

Algorithm **Blocks_PP2a_Numbering Inversion**(g)	Algorithm **Blocks_ PP2a_ Numbering**(id)
Input: block number s Output: block identifier id= [a,b,c]	Input: block identifier id = [a,b,c] Output: block number s

if s = 0: return [0, 0, 1]
if s - 1 < n: return [0, 1, s - 1]
for c in range(n):
  for b in range(n):
  if s = (1 + n + c + bn):
  return [1, b, c]
	if [a, b, c] = [0, 0, 1]:
	  return 0
	if [a, b, c] = [0, 1, c]:
	  return c + 1
	if [a, b, c] = [1, b, c]:
	  return 1 + n + c + bn

Table 28. Algorithms for PP2-block (based on the identifiers balancing) elements numbering and inverse numbering

Algorithm **Elements_PP2a_ Numbering Inversion**(g)	Algorithm **Elements_ PP2a_ Numbering**(id)
Input: element number g Output: element identifier id= [x,i]	Input: element identifier id = [x,i] Output: element number g

if g = 0: return [1, 0, 0]
if g - 1 < n: return [g - 1, 1, 0]
for x in range(n):
  for y in range(n):
  if g = (1 + n + x + yn):
  return [x, y, 1]
	if [x, y, z] = [1, 0, 0]: return 0
	if [x, y, z] = [x, 1, 0]: return x + 1
	if [x, y, z] = [x, y, 1]:
	  return 1 + n + x + yn

Example 1. According to Algorithm **Block_ STS_Skolem(*id*)** (see Table 3)

no more than two arithmetic operations in the ring Z_v are executed regardless of the order of the block design. Hence the asymptotic complexity is equal to the asymptotic complexity of addition in this case. The asymptotic complexity of Algorithm **Dual_Block_PG**3 (*id*) (Табле 6) equals the asymptotic complexity of multiplication in the ring Z_n (two arithmetic operations in this ring). For the algorithms of intersection blocks or dual blocks such as **Dual_Blocks_ PG2_Intersection**($id1$,$id2$) one inversion operation in F_n is executed hence its complexity determines the complexity of this algorithm.

Table 29. Algorithms for PG3-blocks (based on the identifiers balancing) numbering and inverse numbering

Algorithm **Blocks_PG3a_Numbering Inversion**(g)	Algorithm **Blocks_ PG3a_ Numbering**(id)
Input: block number s Output: block identifier id= [a,b,c,d]	Input: block identifier id = [a,b,c,d] Output: block number s

if s = 0:
  return [0, 0, 0, 1]
if s - 1 < n:
  return [0, 0, 1, s – 1]
if s - 1 < n^2 + n:
  for d in range(n):
  for c in range(n):
  if s = (1 + n + d + cn):
  return [0, 1, c, d]
for b in range(n):
  for d in range(n):
  for c in range(n):
  if s = (1+n+n^2+d+cn+bn^2):
  return [1, b, c, d]
	match [a,b,c,d]:
	  case [0,0,0,1]: return 0
	  case [0,0,1,d]: return 1+d
	  case [0,1,c,d]: return 1+n+d +cn
	  case [1,b,c,d]:
	  return 1+n+n^2+d+nc+n^2b

Table 30. Algorithms for PG3-block (based on the identifiers balancing) elements numbering and inverse numbering

Algorithm **Elements_PG3a_ Numbering Inversion**(g)	Algorithm **Elements_ PG3a_ Numbering**(id)
Input: element number g Output: element identifier id= [x,i]	Input: element identifier id = [x,y] Output: element number g

if g = 0:
  return [1, 0, 0, 0]
if g - 1 < n:
  return [g – 1, 1, 0, 0]
if g - 1 < n^2 + n:
  for w in range(n):
  for x in range(n):
  if s = (1 + n + w + xn):
  return [w, x, 1, 0]
for w in range(n):
  for x in range(n):
  for y in range(n):
  if s = (1+n+n^2+w+xn+ yn^2):
  return [w, x, y, 1]
	match [w,x,y,z]:
	  case [1,0,0,0]: return 0
	  case [w,1,0,0]: return 1+ w
	  case [w,x,1,0]: return 1+n+w+xn
	  case [w,x,y,1]: return 1+n+n^2+nx+ n^2y

Table 31. Algorithms for TD2-blocks numbering and inverse numbering

Algorithm **Blocks_TD2_ Numbering Inversion**(g)	Algorithm **Blocks_TD2_ Numbering**(id)
Input: block number s Output: block identifier id= [x,i]	Input: block identifier id = [a,b] Output: block number s
if s = 0: return [0, 0] if s - 1 < p: return [x^{s-1}, 0] else: $m0$ = s mod p $m1$ = (s / p) mod p return [x^{m0}, x^{m1}]	match [a,a]: case [0,0]: return 0 case [a,0]: return ind(a) case [0,b]: return n*ind(b) case _: return ind(a)+n*ind(b)

Table 32. Algorithms for TD2-block and TD#-block elements numbering and inverse numbering

Algorithms **Elements_ TD2_ Numbering_Inversion**(g) **Elements_ TD3_ Numbering_Inversion**(g)	Algorithm **ElementsTD2_ Numbering**(id) **ElementsTD3_ Numbering**(id)
Input: element number g Output: element identifier id= [x,i]	Input: element identifier id = [x,i] Output: element number g

if g = 0:
return [0, 0]
if g - 1 < p:
return [x^{g-1}, 0]
else:
$m0$ = g mod n
$m1$ = (g / n) mod n
return [x^{m0}, x^{m1}]
	match [x,y]:
	case [0,0]: return 0
	case [x,0]: return ind(x)
	case [0,y]: return n*ind(y)
	case _: return ind(x)+n*ind(y)

Table 33. Algorithms for TD3-blocks numbering and inverse numbering

Algorithm **Blocks_TD3_Numbering Inversion**(g)	Algorithm **Blocks_TD3_ Numbering**(id)
Input: block number s Output: block identifier id= [x,i]	Input: block identifier id = [a,b] Output: block number s

if s = 0:
return [0, 0, 0]
if s - 1 < n:
return [x^{s-1}, 0, 0]
if s - 1 < n^2 + n:
$m0$ = s mod n
$m1$ = (s / n) mod n
return [x^{m0}, x^{m1}, 0]
else:
$m0$ = s mod n
m = s / n
$m1$ = m mod n
$m2$ = (m / n) mod n
return [x^{m0}, x^{m1}, x^{m2}]
	match [a,b,c]:
	case [0,0,0]: return 0
	case [a,0,0]: return ind(a)
	case [0,b,0]: return n*ind(b)
	case [0,0,c]: return n^2*ind(c)
	case [a,b,0]: return ind(a)+n*ind(b)
	case [a,0,c]: return ind(a)+n^2*ind(c)
	case [0,b,c]: return n*ind(b)+n^2*ind(c)
	case _: return ind(a)+n*ind(b)+n^2*ind(c)

Table 34. Algorithms for TD3-block in numerical notation construction

Algorithm **Block_TD2N**(g)	Algorithm **Dual_Block TD2N**(e)
Input: block number g Output: list of block elements numbers	Input: element number e Output: list of numbers of blocks containing given element
id=**Blocks_TD2_Numbering_ Inversion**(s) B=**Block_TD2**(id) BN=[] for it in B: BN.append (**ElementsTD2Numbering** (it)) return BN	id= **Elenets_TD2_Numbering_ Inversion** (e) D= **Dual_Block_TD2**(id) DN=[] for it in D:DN.append (**BlocksTD2Numbering** (it)) return DN

5.2 Same Examples of Algorithms Implementation

Below there are given examples of concerning Steiner triples assuming the global parameter t = 2 is set for both types of STS.

Table 35. Algorithms for combined PP2 block and dual block computing

Algorithm **CPP2()**	Algorithm **CDPP2()**
Input: global options Output: list of pairs [block number, list of block element numbers]	Input: global options Output: list of pairs [element number, list of block numbers]
L=[] $v=n^2+n+1$ for i in range $(0,v)$: L.append[i, **Block_PP2N**(i)]	L=[] $v=n^2+n+1$ for i in range $(0,v)$: L.append[i, **DualBlock_PP2N**(i)]

Example 2. The Blocks and Dual blocks of Bose STS computing.

STS_Bose

[0: [0, 5, 10], 1: [1, 6, 11],
2: [2, 7, 12], 3: [3, 8, 13],
4: [4, 9, 14], 5: [0, 1, 8],
6: [0, 2, 6], 7: [1, 2, 9],
8: [0, 3, 9], 9: [1, 3, 7],
10: [2, 3, 5], 11: [0, 4, 7],
12: [1, 4, 5], 13: [2, 4, 8],
14: [3, 4, 6], 15: [5, 6, 13],
16: [5, 7, 11], 17: [6, 7, 14],
18: [5, 8, 14], 19: [6, 8, 12],
20: [7, 8, 10], 21: [5, 9, 12],
22: [6, 9, 10], 23: [7, 9, 13],
24: [8, 9, 11], 25: [10, 11, 3],
26: [10, 12, 1], 27: [11, 12, 4],
28: [10, 13, 4], 29: [11, 13, 2],
30: [12, 13, 0], 31: [10, 14, 2],
32: [11, 14, 0], 33: [12, 14, 3],
34: [13, 14, 1]]

DSTS_Bose

[0: [0, 5, 6, 8, 11, 32, 30],
1: [1, 7, 9, 12, 5, 26, 34],
2: [2, 10, 13, 6, 7, 29, 31],
3: [3, 14, 8, 9, 10, 33, 25],
4: [4, 11, 12, 13, 14, 28, 27],
5: [0, 15, 16, 18, 21, 12, 10],
6: [1, 17, 19, 22, 15, 6, 14],
7: [2, 20, 23, 16, 17, 9, 11],
8: [3, 24, 18, 19, 20, 13, 5],
9: [4, 21, 22, 23, 24, 8, 7],
10: [0, 25, 26, 28, 31, 22, 20],
11: [1, 27, 29, 32, 25, 16, 24],
12: [2, 30, 33, 26, 27, 19, 21],
13: [3, 34, 28, 29, 30, 23, 15],
14: [4, 31, 32, 33, 34, 18, 17]]

Example 3. The Blocks and Dual blocks of Skolem STS computing

STSSkolem

[0: [0, 4, 8], 1: [1, 5, 9],
2: [0, 1, 6], 3: [0, 2, 5],
4: [1, 2, 7], 5: [0, 3, 7],
6: [1, 3, 4], 7: [2, 3, 6],
8: [4, 5, 10], 9: [4, 6, 9],
10: [5, 6, 11], 11: [4, 7, 11],
12: [5, 7, 8], 13: [6, 7, 10],
14: [8, 9, 2], 15: [8, 10, 1],
16: [9, 10, 3], 17: [8, 11, 3],
18: [9, 11, 0], 19: [10, 11, 2],
20: [12, 2, 4], 21: [12, 6, 8],
22: [12, 10, 0], 23: [12, 3, 5],
24: [12, 7, 9], 25: [12, 11, 1]]

DSTSSkolem

[0: [0, 2, 3, 5, 18, 22],
1: [1, 4, 6, 2, 15, 25],
2: [7, 3, 4, 19, 14, 20],
3: [5, 6, 7, 17, 16, 23],
4: [0, 8, 9, 11, 6, 20],
5: [1, 10, 12, 8, 3, 23],
6: [13, 9, 10, 7, 2, 21],
7: [11, 12, 13, 5, 4, 24],
8: [0, 14, 15, 17, 12, 21],
9: [1, 16, 18, 14, 9, 24],
10: [19, 15, 16, 13, 8, 22],
11: [17, 18, 19, 11, 10, 25],
12: [21, 22, 20, 24, 25, 23]]

Example 4. The separate blocks of Bose STS and Skolem STS computing.

BlockSTSBoseN(3)
[3, 8, 13]

BlockSTSSkolemN(10)
[5, 6, 11]

Example 4.

Example 5. The separate blocks of Bose STS and Skolem STS computing.

DualBlockSTSBoseN(6)
[1, 17, 19, 22, 15, 6, 14]

DualBlockSTSSkolemN(3)
[5, 6, 7, 17, 16, 23]

Example 6. Intersection of two blocks Bose STS or Skolem STS computing.

BlockSTSBoseIntersectionN(8,10)
[3]

BlockSTSSkolemIntersectionN(7,13)
[6]

Example 7. Intersection of two dual blocks Bose STS or Skolem STS computing.

DualBlockSTSBoseIntersectionN(12,14).
[33]

DualBlockSTSSkolemIntersectionN(0,4)
[0]

Example 8. Deriving of the list of blocks interseting with each of two nonintersecting blocks of Bose STS or Skolem STS.

AdjasentSTSBoseIntersectionN(0,1)
[11, 8, 32, 12, 21, 18, 28, 22, 31]

AdjasentSTSSkolemIntersectionN(6,10)
[1, 2, 25, 23, 7, 17, 8, 9, 11]

6 Conclusion

Algorithms for calculating individual blocks and dual blocks of CBD and CCBD, which determine the composition of subsystems, as well as intersections of blocks and dual blocks, and in the case of the absence of intersection, algorithms for calculating blocks that intersect with each of two non-intersecting blocks that define connections between subsystems have been represented. For practical application, on their basis, the corresponding algorithms in numerical notation are obtained. In this case, algorithms for numbering and numbering inversion of algebraic identifiers of elements and blocks are used. The complexity of the above algorithms is estimated as complexity of composite algebraic operations (comparisons, multiplications, additions, subtractions, divisions, or square roots) for each list identifier at the output of the algorithm: a limited number of algebraic operations is required to calculate one identifier.

References

1. Solari Esfehani, N., Haj Seyyed Javadi, H.: A survey of key pre-distribution schemes based on combinatorial designs for resource-constrained devices in the IoT network. Wireless Netw. **27**(4), 3025–3052 (2021). https://doi.org/10.1007/s11276-021-02629-8
2. Tarannikov, Y.V.: Combinatorial Properties of discrete structures and applications to cryptology. MCCME, Moscow (2014). — ISBN 978-5-4439-2045-0. (In Russian)
3. Parkhomenko, P.: Algorithmizing design of a class of combinatorial block diagrams. Autom. Remote. Control. **77**(7), 1216–1224 (2016). https://doi.org/10.1134/S0005117916070080

4. Mozharov, G.P.: Fault-tolerant computer networks constructed on the basis of combinatorial block designs, Vestn. Mosk. Gos. Tekh. Univ. im. N.E. Baumana, Priborostr [Herald of the Bauman Moscow State Tech. Univ., Instrum. Eng.], vol. 6, pp. 41–53 (2016). https://doi.org/10.18698/0236-3933-2016-6-41-53. (In Russian)
5. Bahrami, P.N., Javadi H.H.S., Dargahi T., Dehghantanha, A., Choo, K.K.R.: A hierarchical key pre-distribution scheme for fog networks. Concurrency and Computation: Practice and Experience **31**(22) (2019). https://doi.org/10.1002/cpe.4776
6. El Gholami, K., Yassine, M.Y., Fatani, I.: The IEEE 802.15.4 Standard in Industrial Applications: a survey. J Theor. Appl. Inform. Technol. **99**(15), 1–17 (2021)
7. Kim, S.H., Chong, P.K., Kim, T.: Performance study of routing protocols in ZigBee wire-less mesh networks. Wireless Pers. Commun. **95**, 1829–1853 (2017)
8. Kim, T., Kim, S.H., Yang, J., Yoo, S.-E., Kim, D.: Neighbor table based shortcut tree routing in ZigBee wireless networks. IEEE Trans. Parallel Distrib. Syst. **25**(3), 706–716 (2014)
9. Gomez, C., Paradells, J.: Wireless home automation networks: a survey of architectures and technologies. IEEE Commun. Mag. **48**(6), 92–101 (2010)
10. Frolov, A., Kochetova, N.: Networks with the structure of a combined combinatorial block design. In: International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), pp. 1–6 (2021). https://doi.org/10.1109/ICECCME52200.2021.9591136
11. Frolov, A., Kochetova, N.: Scaling networks and capturing keys using combined systems of sets. In: Kravets, A., Bolshakov, A., Shcherbakov, M. (eds.) Society 5.0: Cyber-Solutions for Human-centric technologies. Studies in Systems, Decision and Con-trol, vol 437. Springer, Cham. (2023). https://doi.org/10.1007/978-3-031-35875-3_21
12. Lee, J., Stinson, D.R.: On the construction of practical key predistribution schemes for distributed sensor networks using combinatorial designs. ACM Trans. Inform. Syst. Secur. **11**(2), 1–35 (2008). https://doi.org/10.1145/1330332.1330333
13. Ruj, S., Roy, B.: Key predistribution using combinatorial designs for grid-group deployment scheme in wireless sensor networks. ACM Trans. Sens. Netw. **6**(1), (2009). https://doi.org/10.1145/1653760.1653764
14. Kumar, A., Pais, A.R.: A new combinatorial design based key pre-distribution scheme for wireless sensor networks. J. Ambient. Intell. Humaniz. Comput. **10**(6), 2401–2416 (2019)
15. Masaeli, N., Javadi, H.H., Erfani, H.: Key pre-distribution scheme based on transversal design in large mobile fog networks with multi-clouds. J. Inf. Secur. Appl. **54** (2020). https://doi.org/10.1016/j.jisa.2020.102519
16. Urivskiy, A.: Key Predistribution Scheme Using Affine Planes and Blom's Scheme. In Proceedings of the International Mathematics Conference "50 Years OF IITP" (2011). — ISBN 978-5-901158-15-9.
17. Frolov, A., Kochetova, N., Klyagin, A.: Scaling networks and capturing keys using combined combinatorial block designs. In: Silhavy, R., Silhavy, P. (eds.) Networks and Systems in Cybernetics: Proceedings of 12th Computer Science On-line Conference 2023, Volume 2, pp. 269–289. Springer International Publishing, Cham (2023). https://doi.org/10.1007/978-3-031-35317-8_26
18. Frolov, A., Klyagin, A., Kochetova, N., Temnikov, D.: On the synthesis of systems possessing the structure of some combinatorial designs. J. Comput. Syst. Sci. Int. **60**(4), 585–594 (2021). https://doi.org/10.1134/S1064230721040067

Spectral Method of Analysis and Complex Processing of Video Information from Multi-Zone Images

Valery V. Vasilevskiy(✉)

Moscow Aviation Institute (National Research University), 4, Volokolamskoe Highway, Moscow 125993, Russia
echinops777@rambler.ru

Abstract. The problem of increasing the efficiency of monitoring of low-contrast objects under conditions of a priori uncertainty of the interference-signal environment based on the technology of complex multi-channel processing of spectro-zonal information of multi-zone images obtained by receiving devices is investigated. The stochastic filtering solution is reduced to a simpler deterministic problem of estimating the a posteriori density of the observational data distribution based on the spectral method. We construct an image model based on the spectral characteristics of the random anomaly brightness field and the state vector, represented as a series by functions of the basis system. A description of the adaptive computational algorithm for the complex processing and analysis of the anomaly state of the image based on the approximate solution of the system of linear algebraic equations by the spectral method is given. #COMESYSO1120.

Keywords: Multi-Zone Image · Observation System Model · Distribution Density Function · Spectral Method · Spectral Characteristic · Basis Function System · Adaptive Estimation Algorithm

1 Introduction. Problem Statement

Observation systems that implement processes of reception and processing of multiple-zone images (MZI) with a random anomaly structure are mathematical models of stochastic control systems for aerospace monitoring devices of different types [1, 2].

Since the analytical solution for such systems can be found only for particular cases, the most widespread for practical tasks are approximate methods, which can be divided into two groups.

The first group includes methods based on direct modeling of the observation system under the influence of random external perturbations and the implementation of following statistical processing of results of information analysis.

The second group is characterized by the transition from the original problem to the deterministic one associated with the solution of equations for the probability density or moment characteristics of MZI formed in the observation system.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 330–335, 2024.
https://doi.org/10.1007/978-3-031-53552-9_28

In this paper we consider the solution of the problem of complex multichannel processing of spectrozonal images obtained by receivers of the corresponding range as the problem of analysis of stochastic observing system and estimation of the detected anomalies state based on spectral method.

Here, we will refer to the class of anomalies the observation objects, the state vector (structure and coordinates) of which is a random function of observation.

The approach used is based on the transition to the deterministic analysis problem based on the use of the spectral method in the construction of the observation system model and the adaptive computational method.

2 System State Model

Model of observation system contains $s = 1, \ldots, S$ spectral channels of processing MZI spectrozonal information [1].

The flow of formed MZI frames is represented as a random field - a set $U\left\{u_{\bar{j}},\ \bar{j} \in J\right\}$ of random variables of the brightness function u_j, changing on multidimensional lattice in discrete space and time $J_t = \left\{\bar{j} = (j_1, j_2, j_3);\ j_k = \overline{1, M_l},\ k = 1, 2, 3;\quad t \in [t_0, T]\right\}$.

The random process of changing the brightness function $U\left\{u_{\bar{j}},\ \bar{j} \in J\right\}$ when detecting an anomaly in the current MZI frame can be characterized by the observation state vector $Z(t) = z(t, u)$ as a function of time t and data $u = u(x, y)$ defined on the set $L([t_0, T] \times \Omega)$, $\Omega = \left\{u(x_i, y_j) :\ x_i = \overline{1, m},\ y_j = \overline{1, n}\,; m = n = a\right\}$.

It is required at moments of discrete processing time $t = t_0 + T$ of each current MZI frame to obtain a state estimation vector $\hat{Z}(t)$ in the form of a functional [1, 3] based on the results of all measurements $U_{t_0}^T = \{u(x, y),\ t_0 \le t \le T\}$:

$$\hat{Z}(t) = \psi\left(t, U_{t_0}^T\right) \tag{1}$$

where $\psi\left(t, U_{t_0}^T\right)$ is the estimating functional of MZI spectrozonal information measurements.

This expression defines the input-output relation of the synthesized *statistical nonlinear filter* of the observation system, which evaluates the state vector $z(t, u)$ of the detected MZI anomaly.

Decision of this complex stochastic problem can be reduced to a simpler deterministic problem of finding a partial minimum of the *loss function*:

$$\psi\left(t, U_{t_0}^T\right) = \arg\min_{\psi \in R^n} J\left(t, U_{t_0}^T, \psi\right) \tag{2}$$

where $J\left(t, U_{t_0}^T, \psi\right) = \int\limits_{R^n} (u - \psi)^T C(t)(u - \psi) p\left(t, z|U_{t_0}^T\right) dx$ is the conditional risk function, expressed through posterior density function $p\left(t, z|U_{t_0}^T\right)$; $C(t) > 0$ is the symmetrical positively determined matrix of weight coefficients; $p\left(t,\ z \,\middle|\, U_{t_0}^t\right)$ is the unknown posterior distribution density function of the state vector $Z(t)$ of the detected anomaly.

The distribution density function $p(t, z)$ of the observation system state vector $z(t, u)$ satisfies the system of generalized Fokker-Planck-Kolmgler equations, the solution of which can be constructed using the *spectral method* [5].

The spectral characteristic of the anomaly state vector function $z(t, u)$ on MZI will be called a $(n + 1)$-dimensional matrix $Z(n + 1, 0)$, the elements of which are the coefficients of expansion of this function in series of orthonormalized functions of the basis system $\{e_{i0\ldots in}(t, u)\}_{i_0,\ldots,i_n=0}^{\infty}$ [3, 4]:

$$z_{i0\ldots in} = \int_{t0}^{T} \int_{\Omega} e_{i_0,\ldots i_n}(t, u)\, z(t, u)\, dudt,\ i_0, \ldots, i_n = 0, 1, 2, \ldots \tag{3}$$

The spectral characteristic of the initial state function $z(u)$ of the system at the moment of time $t = t_0$ is called the n-dimensional matrix $Z(n, 0)$ with the elements

$$z_{i1\ldots in} = \int_{\Omega} \chi_{i1\ldots in}(x) z(x) dy,\ i_1, \ldots, i_n = 0, 1, 2, \ldots \tag{4}$$

where $e_{i_0\ldots i_n}(t, u) = q_{i_0}(t)\chi_{i_1\ldots i_n}(u)$, $i_0, \ldots, i_n = 0, 1, 2, \ldots$ is the orthonormalized basis of the space $L([t_0, T] \times \Omega)$ of time functions and the state vector of the system.

The spectral characteristics of the distribution density functions of the state vector $p(t, z)$ and $p_0(z)$ are represented in the form of matrices $\Phi(n + 1, 0)$ and $\Phi(n, 0)$, respectively, and satisfy the system of equations of *generalized characteristic functions* [5]. An approximate solution of the problem by the spectral method can be obtained using a system of linear algebraic equations of the form [3, 5]:

$$p(t, z) = \sum_{i_0=0}^{L-1} \cdots \sum_{i_n=0}^{L-1} \phi_{i_1\ldots i_n} e_{i_1\ldots i_n}(t, z) \tag{5}$$

where the matrix $\Phi^{(l)}(n + 1, 0)$ elements $\phi_{i_0\ldots i_n}$ are coefficients of expansion of the required non-normalized distribution densities $p(t, z)$ in series on the basis system functions $\left\{e_{i_0\ldots i_n}(t, z)\right\}_{i_0,\ldots,i_n=0}^{\infty}$; L is the order of spectral characteristics truncation.

The state vector $z(t, u)$ for a stationary (finite) time interval $\theta = [t_0, T]$ of MZI frame processing can be represented as a series on the basis system functions $\{e(i_1, u)e(i_2, u)\}_{i_1,i_2=0}^{L}$ on the set Ω:

$$z(t, u) = \sum_{i_1 i}^{L} \sum_{i_2}^{L} z_{i_1 i_2} \cdot e(i_1, i_2, u),\ u \in \Omega \tag{6}$$

where the decomposition coefficients z_{i_1,i_2} are calculated by the formula:

$$z_{i_1 i_2} = (e(i_1, i_2, u), z(t, u))_{L(\Omega)},\ i_1, i_2 = 0, 1, 2, \ldots \tag{7}$$

As a basis system of state vector functions orthonormalized on the set Ω, we use a system of trigonometric functions of the form

$$e(i, u) = \begin{cases} \sqrt{\frac{1}{2a}},\ i = 0, \\ \sqrt{\frac{1}{a}} \cos\left(\frac{i\pi}{2a}\right) u,\ i = 2k, \\ \sqrt{\frac{1}{a}} \sin\left(\frac{(i+1)\pi}{2a} u\right),\ i = 2k - 1. \end{cases} \tag{8}$$

Thus, the solution of the problem of representation of the state vector function $z(t, u) \in \Omega$ of MZI in the form of a series by functions of the selected basis system consists in calculation of decomposition coefficients z_{i_1,i_2} and use of expression (6).

3 Computational Algorithm

The basis of complex processing and synthesis of a single image for solving the problems of analysis of detected anomalies is an adaptive estimation in spectral channels of two-dimensional data of a random field of multi-zone images [1, 6].

Taking into account the constructed observation model, the adaptive algorithm of complex processing and estimation of MZI based on the approximate solution of the system of linear algebraic equations by spectral method includes the following sequence of computational procedures.

Step 1. Form a basis system of functions of the argument of time and state vector $\left\{e_{i_0 \ldots i_m}(t, z)\right\}_{i_0,\ldots,i_n=0}^{\infty}$, using expression (8), Select the order of truncation of spectral characteristics L.

Step 2. Represent the function of the state vector $z(t, u)$ in the form of series (6) by functions of the basis system $\{e(i_1, x_1)e(i_2, x_2)\}_{i_1,i_2=0}^{L}$ on the set $L([t_0, T] \times \Omega)$, $\Omega = \left\{u(x_i, y_j) : x_i = \overline{1, m},\ y_j = \overline{1, n}\,; m = n = a\right\}$.

Step 3. Find the solution of the problem by spectral method using the system of linear algebraic equations, represent the distribution densities $p(t, z)$ of the anomaly MZI state vector for spectral channels $s = 1, \ldots, S$ in the form of expression (5).

Step 4. Obtain an estimate of the functional state vector $\hat{Z}(t)$, taking into account the loss function criterion of the form (2) for each spectral channel, construct the structure of the detected anomaly $L([t_0, T] \times \Omega)$.

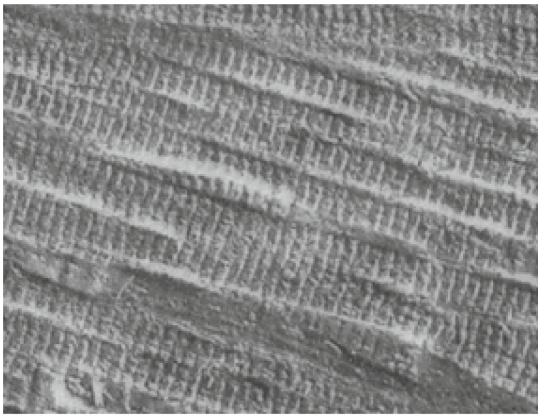
In the procedure of automatic complex estimation of MZI anomaly state vector, the corresponding frames of spectrozonal video fragments are combined in the synthesis of a single grayscale or color image.

The results of using the spectral adaptive algorithm for multispectral image processing on the test task of synthesizing a single image and detection of local terrain inhomogeneities of the selected image fragment are shown in Fig. 1.

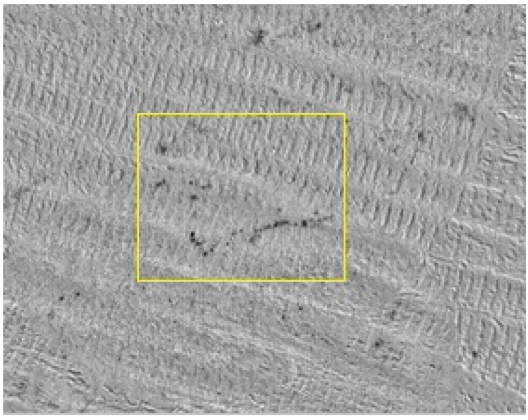
On the original spectrozonal image, the anomaly (low-contrast object - spots of oil products) is little noticeable, on the synthesized halftone unified image, the observed image highlighted by a frame is detected with greater reliability.

From the analysis of the results of testing the algorithm on model problems, we can conclude that the visibility of the anomaly structure in the synthesized single image can be increased compared with the original spectrozonal images.

The efficiency of the MZI processing and analysis algorithm in assessing the state of the detected anomaly depends on the signal-to-noise ratio of the original spectrozonal images, the type of object and background textures, the choice of orthonormalized system of basis functions and the order of truncation of spectral characteristics.



a)



b)

Fig. 1. Spectral method of analysis: a) spectrozonal image; b) synthesized image.

4 Conclusion

The problem of increasing the efficiency of low-contrast objects monitoring under a priori uncertainty of the interference-signal environment based on the technology of complex multi-channel processing of spectrozonal information of multi-zone images obtained by receiving devices has been investigated.

The problem of stochastic filtering of information of multi-zone images can be solved by setting a penalty function and switching to a simpler deterministic problem of estimating the a posteriori density of data distribution of the detected MZI anomaly.

To build an observation model based on the proposed approach, it is necessary to represent the spectral characteristics of the random brightness field and the state vector of the detected anomaly as a series by orthonormalized functions of the basis system.

Adaptive computational algorithm of multichannel complex processing and analysis of MZI when evaluating the structure of the detected anomaly can be constructed based on an approximate solution of the system of linear algebraic equations by spectral method, the efficiency of which depends on the quality of the obtained initial spectrozonal images.

References

1. Vasilevskiy, V.: Deterministic-probabilistic image model in technologies of aerospace monitoring. AIP Conf. Proc. **2700**, 020025 (2023). https://doi.org/10.1063/5.0125838
2. Fukunaga, K.: Introduction to the Statistical Theory of Image Recognition. Nauka, Moscow (1979)
3. Gurman, V.I.: Principle of Expansion in Control Problems. Fizmatlit, Moscow (1997)
4. Ditkin, V.A., Prudnikov, N.P.: Integral Transformations and Operational Calculus. Fizmatgiz, Moscow (1974)
5. Pugachev, V.S., Sinitsyn, I.N.: Theory of Stochastic Systems. Logos, Moscow (2000)
6. Fomin, V.N.: Recurrent Estimation and Adaptive Filtration. Nauka, Moscow (1984)

Cyber-Physical Control System for Personal Protective Equipment Against Infectious Diseases Transmitted by Airborne Droplets

Mikhail Golosovskiy[1], Alexey Bogomolov[2(✉)], Eugene Larkin[3], and Tatiana Akimenko[3]

[1] St. Petersburg Federal Research Center of the Russian Academy of Sciences, 39, 14th Line Vasilyevsky Island, St. Petersburg 199178, Russia
[2] Federal Research Center «Computer Science and Control» of the Russian Academy of Sciences, 44 Building 2 Vavylova Street, Moscow 119333, Russia
a.v.bogomolov@gmail.com
[3] Tula State University, 92, Lenin Avenue, Tula 300012, Russia

Abstract. An approach to modeling the fault tolerance of components of a cyber-physical control system for personal protective equipment against infectious diseases transmitted by airborne droplets based on the theory of semi-Markov processes is presented. At the first stage of the implementation of the approach, the life cycles of each component of the control system are considered with the construction of a conventional semi-Markov model. At the second stage, the model of competition between the components of the control system is considered in terms of time between failures. To simplify modeling, a digital approach to competition analysis has been developed. A digital model of the fault tolerance of all components of the control system is obtained using the sampling density, which is described by the uptime of the components. It is shown that the digital approach to the analysis of competition makes it quite simple to calculate the service life of the components of the control system and the control system as a whole. At the same time, the number of samples on the histogram, represented by a lifetime unit, allows you to control the accuracy and computational complexity of calculating the reliability characteristics of the control system.

Keywords: digital control system · cyber-physical system · semi-Markov process · competition model · control system reliability · control system fault tolerance

1 Introduction

Studies of the quality of personal protective equipment against infectious diseases transmitted by airborne droplets (PPE), conducted in various countries of the world, show that the use of most PPE samples leads to a significant increase in the severity of the labor process [1, 2]. This is due to excessive restriction of movements, suboptimal microclimate in the underwear space, obstacles to communication between personnel, lack of PPE status indicators, etc. [3–5].

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 336–342, 2024.
https://doi.org/10.1007/978-3-031-53552-9_29

Obviously, most of these shortcomings can be eliminated by improving the design and ergonomic characteristics of PPE. The priorities of such improvements are currently associated with the introduction of technologies of the fourth industrial revolution - an information technology concept that implies the integration of computing resources into physical processes.

When implementing PPE based on such a concept, sensors, equipment and information systems will be connected to interact with each other using Internet protocols, providing the ability to predict, self-adjust and adapt PPE to the user's state and to changes in the external environment in real time [6–9].

Ensuring the adjustment of the synthesized adaptive control algorithms for PPE can be implemented using machine learning, and taking into account the well-known characteristics of the control object, in addition to using classical controllers, it is promising to use methods of robust feedback linearization, as well as robust and indirect adaptive control based on the pacification principle [10].

To increase the reliability of the control system, it is necessary to provide redundancy of sensors and computing modules (computing components). Correct redundancy planning presupposes the presence of models of the failure/recovery process of both individual components of the control system and the control system as a whole.

Analysis of the subject area shows that the most effective approach to solving such problems is based on the application of the theory of Markov or semi-Markov processes [11, 12].

However, when modeling redundant structures, a competition effect arises, to minimize the effects of which semi-Markov models are often replaced by coarser Markov models, which noticeably reduces the modeling accuracy [11, 13, 14]. Therefore, as a compromise, discrete semi-Markov models are used instead of Markov models to describe competition in fault-tolerant control systems. This provides not only the ability to control the accuracy of control actions, but also the ability to vary them in accordance with the required control reliability.

2 Methods

The components of the cyber-physical control system of PPE and, in which the principle of fault tolerance is implemented, can be considered as M units operating in parallel [15–17]. Faults/recoveries in such units develop in parallel, therefore, to describe the reliability of the functioning of the entire system, its representation in the form of an M-parallel semi-Markov process can be used:

$$\mu = \{\mu_1, \ ..., \ \mu_m, \ ..., \ \mu_M\},$$

where $\mu_m = \{A_m, \ \boldsymbol{h}_m(t)\}$, $1 \le m \le M$, is the ordinary semi-Markov process [18–22], which is characterized with set of structural states $A_m = \left\{a_{0(m)}, \ ..., \ a_{j(m)}, \ ..., \ a_{J(m)}\right\}$ and semi-Markov matrix

$$\boldsymbol{h}_m(t) = \left[h_{j(m),k(m)}(t)\right];$$

t is the time; $a_{0(m)}$ simulates the start of m-th unit exploitation when it surely able to work; $a_{J(m)}$ is the absorbing state, which simulates the fully destroyed unit; $a_{j(m)}$, $1(m) \leq j(m) < J(m)$ simulate other physical states (able to work, short-time failed, under recovering, etc.);

$$\boldsymbol{h}_m(t) = \boldsymbol{p}_m \otimes \boldsymbol{f}_m(t);$$

$\boldsymbol{p}_m = \left[p_{j(m),k(m)}\right]$ and $\boldsymbol{f}_m(t) = \left[f_{j(m),k(m)}(t)\right]$ are $[J(m)+1] \times [J(m)+1]$ stochastic matrix and matrix of pure time densities, correspondingly [23–26].

3 Results and Discussion

When solving problems, we believe that a certain fault-tolerant set of components (elements) of the control system is operational as long as at least one of all M blocks remains "alive". As an example, consider a fault-tolerant cyber-physical control system shown in Fig. 1.

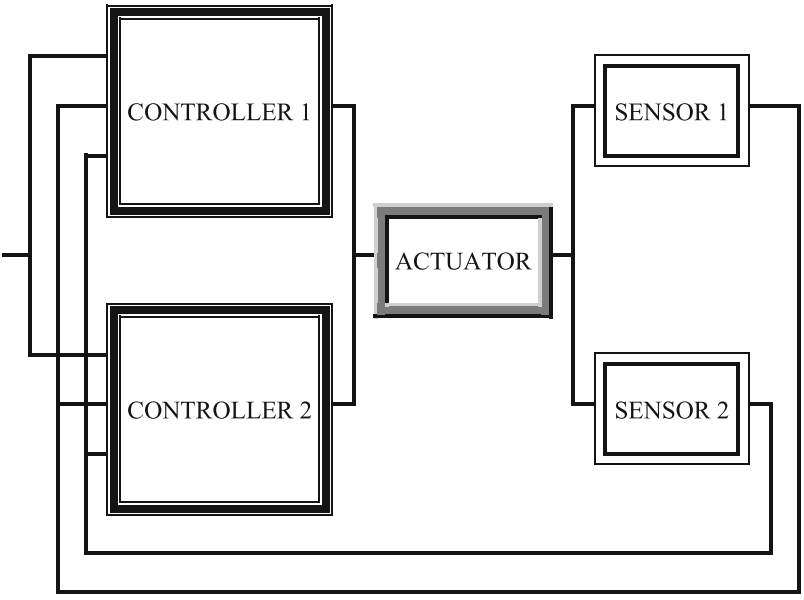


Fig. 1. Fault tolerant actuator control system.

To improve reliability, the system includes a duplicated controller and a sensor connected to the circuit in such a way that both sensors measure the same value, for example, the air temperature in the PPE underwear space, and the measurement results of each sensor are transmitted to both controllers. The electronic components of the cyber-physical control system of the actuators are affected by correlated noise that occurs when the power is turned on/off for the volatile components of the control system. Interference

can cause both short-term malfunctions and serious failure of PPE control system components. The entire system is inoperable only if both sensors and/or both controllers fail.

Structure of the 4-parallel semi-Markov process, which describes fault/recoveries in the system, is shown on the Fig. 2:

$$\mu = \{\mu_1,\ \mu_2,\ \mu_3,\ \mu_4\},$$

where

$$\mu_m = \left\{ \{a_{1(m)},\ a_{2(m)},\ a_{3(m)}\},\ \begin{bmatrix} 0 & \delta(t) & 0 \\ 0 & h_{2(m),2(m)}(t) & h_{2(m),3(m)}(t) \\ 0 & 0 & 0 \end{bmatrix} \right\}, m \in \{1,\ 2,\ 3,\ 4\}.$$

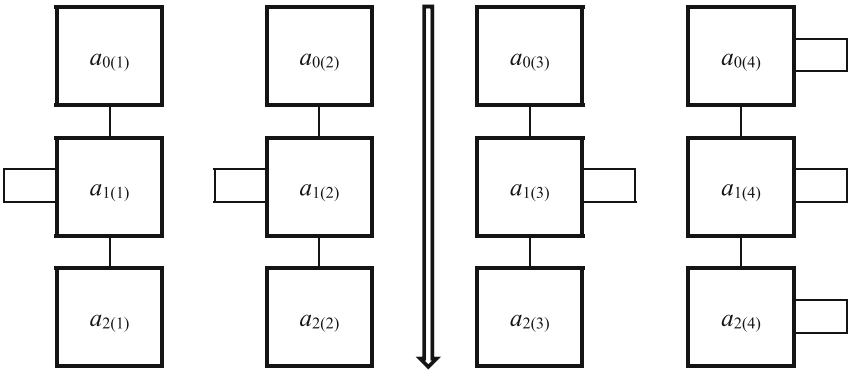


Fig. 2. The parallel semi-Markov process, describing fault/recoveries

In (43) $a_{1(m)}$ and $a_{3(m)}$ are the starting state (start of the system operation) and the absorbing state (equipment hard failure), correspondingly; $\delta(t)$ is the Dirac δ-function; $h_{2(m),i(m)}(t) = p_{2(m),i(m)} f_{2(m),i(m)}(t)$, $i \in \{2,\ 3\}$, is the weighted time density between noise impulses; $f_{2(m),2(m)}(t)$, $f_{2(m),3(m)}(t)$ are pure time densities; $p_{2(m),2(m)}$, $p_{2(m),3(m)}$ are probabilities of a short-term malfunctions and a hard failure, correspondingly;

$$p_{2(m),2(m)} = \begin{cases} 0,8, & \text{when } m \in \{1,\ 2\}; \\ 0,9, & \text{when } m \in \{3,\ 4\}; \end{cases}$$

$$f_{2(m),i(m)}(t) = \begin{cases} 2, & \text{when } |t-1| \le 0,25; \\ 0, & \text{when } |t-1| > 0,25, \end{cases} \quad i \in \{2,\ 3\},\ m \in \{1,\ 2,\ 3,\ 4\}.$$

Sampled time densities, obtained from analysis of semi-Markov processes evolution accordingly, and using the sampling method, described in Sect. 3 above, are shown on the Fig. 3, where dashed lines indicate expectations of lifetime of the controller (T_1, T_2) and the sensor (T_3, T_4).

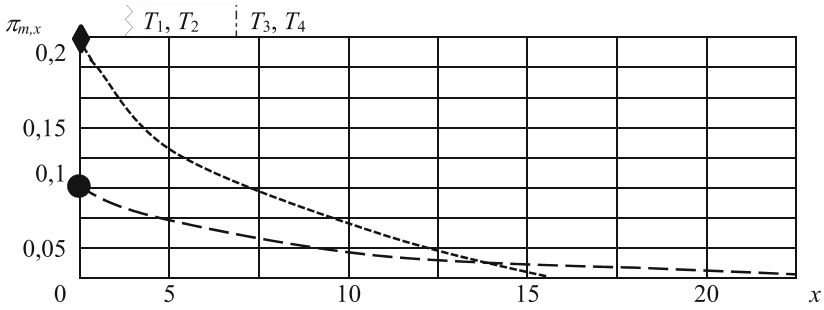


Fig. 3. Sampled time densities (rhomb – $m \in \{1, 2\}$, circle – $m \in \{3, 4\}$)

An analysis of scenarios for the development of failures during the interaction of components of a fault-tolerant system, taking into account the average time to failure and the probability of occurrence of each scenario, calculated using the stated mathematical apparatus, showed that a fault-tolerant structure almost doubles the time of failure-free operation of the control system.

4 Conclusion

The results of the study show that the task of designing a cyber-physical PPE control system should be divided into two stages:

development of conventional semi-Markov models of individual nodes and their transformation into a discrete form;

analysis of a parallel discrete semi-Markov process to obtain the reliability characteristics of components (elements) and the reliability characteristics of the system as a whole.

The proposed approach makes it possible to create a model of a redundant control system with a priori specified accuracy, to develop and implement a method for optimizing a fault-tolerant control system. This will ensure the high reliability of PPE required by the specifics of their operation.

Acknowledgements. The work was supported by the Russian Science Foundation, project no. 22-29-00808.

References

1. Batov, V.E.: Assessment of the functional state of military medical personnel when using personal protective equipment during the covid-19 pandemic. Medico-Biol. Socio-Psychol. Probl. Safety Emerg. Situations **1**, 82–88 (2022)
2. Stewart, C.L., Thornblade, L.W., Diamond, D.J., Fong, Y., Melstrom, L.G.: Personal protective equipment and covid-19: a review for surgeons. Ann. Surg. **272**(2), 132–138 (2020)
3. Ong, J.J.Y., et al.: Headaches associated with personal protective equipment - a cross-sectional study among frontline healthcare workers during Covid-19. Headache **60**(5), 864–877 (2020)

4. Bubeev, Yu.A., Vladimirskiy, B.M., Ushakov, I.B., Usov, V.M., Bogomolov, A.V.: Mathematical modelling of spread covid-19 epidemic for preventive measures to protect life and health of elderly. Bull. South Ural State Univ. Ser. Math. Model. Program. Comput. Softw. **14**(3), 92–98 (2021)
5. De la Torre Bayo, J.J., Martín-Lara, M.A., Calero-Hoces, M., Sánchez-Castillo, P.M., Pula, H.J., Zamorano, M.: Management of used Covid-19 personal protective equipment: a bibliometric analysis and literature review. Appl. Sci. **13**(4), 2377 (2023)
6. Bogomolov, A.: Information technologies of digital adaptive medicine. Inform. Autom. **20**(5), 1153–1181 (2021)
7. Golosovskiy, M.S., Bogomolov, A.V., Balandov, M.E.: Optimized fuzzy inference for Sugeno-type systems. Autom. Doc. Math. Linguist. **56**(5), 237–244 (2022)
8. Golosovskiy, M.S., Bogomolov, A.V., Evtushenko, E.V.: An algorithm for setting Sugeno-type fuzzy inference systems. Autom. Doc. Math. Linguist. **55**(3), 79–88 (2021)
9. Golosovskiy, M.S., Bogomolov, A.V., Terebov, D.S., Evtushenko, E.V.: Algorithm to adjust fuzzy inference system of Mamdani type. Vestn. Yuzhno-Ural. Gos. Univ. Ser. Mat. Mekh. Fiz. **10**(3), 19–29 (2018)
10. Aranovskiy, S., Bobtsov, A., Pyrkin, A.: New results on adaptive systems. Int. J. Adapt. Control Signal Process. **36**(6), 1250–1251 (2022)
11. Spong, M.W., Hutchinson, S., Vidyasagar, M.: Robot Modeling and Control. Wiley, New York (2020)
12. Larkin, E., Privalov, A., Bogomolov, A., Akimenko, T.: Model of digital control system by complex multi-loop objects. In: AIP Conference Proceedings, vol. 2700, p. 030009 (2023)
13. Larkin, E.V., Akimenko, T.A., Bogomolov, A.V.: Self-organizing mobile robots swarm movement control simulation. In: Tan, Y., Shi, Y., Niu, B. (eds.) ICSI 2022. LNCS (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 13345, pp. 56–65. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-09726-3_6
14. Larkin, E., Bogomolov, A., Privalov, A.: Digital control by robot manipulator with improved rigidity. Smart Innov. Syst. Technol. **329**, 45–59 (2023)
15. Ivutin, A.N., Larkin, E.V.: Simulation of concurrent games. Bull. South Ural State Univ. Ser. Math. Modell. Program. Comput. Softw. **8**(2), 43–54 (2015)
16. Larkin, E., Privalov, A., Bogomolov, A., Akimenko, T.: Digital control of continuous production with dry friction at actuators. Smart Innov. Syst. Technol. **232**, 427–436 (2022)
17. Tobin, D., Bogomolov, A., Golosovskiy, M.: Model of organization of software testing for cyber-physical systems. Stud. Syst. Decis. Control **418**, 51–60 (2022)
18. Larkin, E.V., Akimenko, T.A., Bogomolov, A.V.: Modeling the reliability of the onboard equipment of a mobile robot. Izvestiya of Saratov University. Math. Mech. Inform. **21**(3), 390–399 (2021)
19. Larkin, E., Akimenko, T., Bogomolov, A., Krestovnikov, K.: Mathematical model for evaluating fault tolerance of on-board equipment of mobile robot. Smart Innov. Syst. Technol. **187**, 383–393 (2021)
20. Larkin, E., Bogomolov, A., Gorbachev, D., Privalov, A.: About approach of the transactions flow to Poisson one in robot control systems. In: Ronzhin, A., Rigoll, G., Meshcheryakov, R. (eds.) ICR 2017. LNCS (LNAI), vol. 10459, pp. 113–122. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66471-2_13
21. Bauer, H.: Probability Theory. Walter de Gruyter, Berlin (1996)
22. Larkin, E.V., Bogomolov, A.V., Privalov, A.N., Dobrovolsky, N.N.: Discrete model of paired relay-race. Bull. South Ural State Univ. Ser. Math. Model. Program. Comput. Softw. **11**(3), 72–84 (2018)
23. Larkin, E.V., Bogomolov, A.V., Ivutin, A.N.: Statistical estimations of transaction flows in multi-component information systems. In: Distributed Computer and Communication Networks: Control, Computation, Communications (DCCN 2017), pp. 596–603 (2017)

24. Larkin, E.V., Bogomolov, A.V., Privalov, A.N., Dobrovolsky, N.N.: Relay races along a pair of selectable routes. Bull. South Ural State Univ. Ser. Math. Model. Program. Comput. Softw. **11**(1), 15–26 (2018)
25. Larkin, E., Akimenko, T., Bogomolov, A., Sharov, V.: Reliability of robot's controller software. In: Ronzhin, A., Sadigov, A., Meshcheryakov, R. (eds.) ICR 2023. LNCS, vol. 14214, pp. 289–299. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-43111-1_26
26. Larkin, E.V., Akimenko, T.A., Bogomolov, A.V.: The swarm hierarchical control system. In: Tan, Y., Shi, Y., Luo, W. (eds.) ICSI 2023. LNCS (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 13968, pp. 30–39. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-36622-2_3

Concept of Smart Personal Protection Equipment Against Infectious Diseases

Alexey Bogomolov[1(✉)], Eugene Larkin[2], and Tatiana Akimenko[2]

[1] St. Petersburg Federal Research Center of the Russian Academy of Sciences, 39, 14th Line Vasilyevsky Island, St. Petersburg 199178, Russia
a.v.bogomolov@gmail.com

[2] Tula State University, 92, Lenin Avenue, Tula 300012, Russia

Abstract. The article presents the concept of managing the functioning of intelligent personal protective equipment against infectious diseases to ensure the optimal microclimate inside the personal protective equipment when it is used by the user. The controlled characteristics are the temperature and humidity of the air in the space between the personal protective equipment and the user's body. The control is implemented in the form of a two-level hierarchical control system for the components of personal protective equipment, which provides optimal characteristics of the air mixture in real time. The collection of information is carried out using a variety of temperature and humidity sensors built into the personal protective equipment. The information is transmitted to the controller, coupled with a microcompressor, which ensures the formation of an air mixture with a given temperature and humidity, and also provides the ability to adjust the intensity of the output air flow. The controller and microcompressor are also built into the personal protective equipment. The functioning of the controller in terms of the implementation of control actions is based on the theory of semi-Markov processes and the use of von Neumann type controllers.

Keywords: personal protective equipment control · digital control system · temperature regulation · humidity regulation · semi-Markov process · von Neumann type controllers

1 Introduction

Studies of personal protective equipment against infectious diseases transmitted by airborne droplets (PPE) indicate the need to intellectualize the management of their functioning. This is ensured by methods of intelligent control of cyber-physical systems with guaranteed quality and reliability of the system in relation to the design of PPE from infectious diseases transmitted by airborne droplets [1, 2].

Obviously, such methods should ensure the localization and isolation of faults and compensation for their influence in order to maintain the operability of the cyber-physical PPE control system. This is achieved through the use of system reconfiguration methods, adaptive methods of fault-tolerant control of nonlinear systems, active methods of fault-tolerant control based on the apparatus of swarm intelligence [3].

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 343–351, 2024.
https://doi.org/10.1007/978-3-031-53552-9_30

According to this, the efficiency of solving the problem of joint control of the functioning of PPE components increases if the components are considered as a swarm of jointly functioning elements. Swarm management is a more complex task than managing a single component, the need to synchronize the work of various components [4, 5].

The specificity of solving such problems determines the complexity of the structure of the control system, which should be divided into two hierarchical levels; strategic and functional-logical, each of which solves its own range of tasks [6, 7].

The strategic level is represented by the controller-organizer, whose function is to divide the general task solved by the set of PPE components (swarm) as a whole into local tasks solved by groups of components, transfer tasks to the functional-logical level and receive confirmation of the tasks completion. The PPE components that determine the functional-logical level are included in the feedback that controls the current state of the PPE as a whole [8, 9].

The effectiveness of the controlled operation of the PPE components is determined, on the one hand, by the control algorithm embedded in the controller, and on the other hand, by the operation of the von Neumann type controller, as a physical device that interprets the control algorithm "step by step", and spend time on interpretation. In turn, the reaction time of the computer causes delays in the processing and transmission of information in feedback loops [10–12].

Thus, the digital controller, in addition to the algorithmic implementation of the control law [11], contributes to delays in the control process, which, in turn, affects the quality characteristics of the PPE control system [12]. To simulate time delays, it is advisable to use the apparatus of semi-Markov processes [13, 14], since each control algorithm processes random data generated at the outputs of sensors on the state of PPE and the user of PPE, and implements decision-making operators at branch points.

At the strategic level, PPE components (swarm units) compete with each other for the performance of their tasks that affect the achievement of the global goal of PPE functioning – ensuring safe comfortable working conditions for the PPE user [15–17]. Thus, predicting the behavior of PPE components during competition is necessary to develop an effective algorithm for managing a group of components (swarm) built into the organizing controller [18–21].

Methods for designing hierarchical control systems that take into account the real properties of the von Neumann type controller have not been sufficiently developed to date, and there have been no such studies in relation to PPE control, which determines the relevance of the study.

2 Methods

The structure of a two-level hierarchical system for managing PPE components is shown on the Fig. 1.

The strategic level of management is implemented by the central controller (controller-organizer), which sets the formation of global goals for the functioning of the components of the PPE management system. The global goal in the problem under consideration is to ensure the optimal characteristics of the microclimate in the space

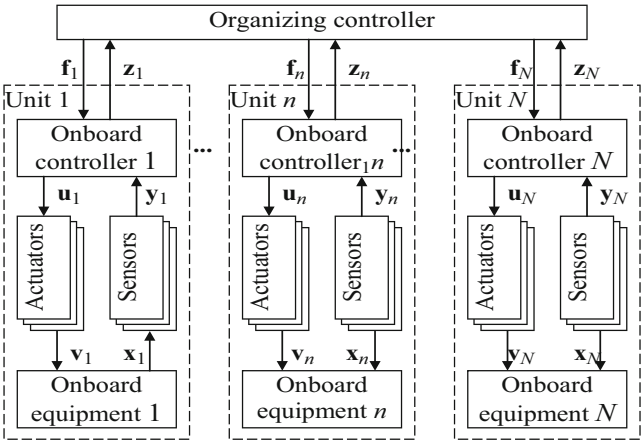


Fig. 1. Flowchart of hierarchical control system

between the PPE and the body of the PPE user. This is achieved with the help of air temperature and humidity sensors connected to the central controller and a mini-compressor designed to absorb air with the issuance of an air mixture with the required characteristics of temperature, humidity and intensity (velocity) of the air flow. The air flow can be given out:

directly into the space between the PPE and the body of the wearer of the PPE;
into plastic tubes built into the PPE with holes directed towards the body of the PPE wearer;
into the plastic tubes embedded in the PPE user's clothing with holes directed towards the user's body;
using a combination of these methods.

Thus, the central controller, in cooperation with the sensors and the compressor, solves the local problems of monitoring and controlling the humidity, temperature and intensity of the air mixture produced by the compressor. In the general case, the number of sensors and compressors can be any, depending on the characteristics of the PPE.

PPE components (swarm) include N sensors and controllers U_i, $i \in \{1 \ldots N\}$, operating under the control of the organizing controller, and the goal can be formulated as providing the specified characteristics of the air environment in the space between the PPE and the body of the PPE user. The organizing controller, receiving information from U_i, about its current Cartesian coordinates z_i, $i \in \{1 \ldots N\}$, must calculate the target vector $\mathbf{f}_n = \left(f_{n,v},\ f_{n,\psi}\right)$, for the n-th component $\mathbf{f}_n = \left(f_{n,v},\ f_{n,\psi}\right)$, where $f_{n,v}$ is the required temperature; $f_{n,\psi}$ - required air humidity.

In the PPE controller, the vector $\mathbf{f}_n$ is converted into a data vector $\mathbf{u}_n = \left(u_{n,l},\ u_{n,r}\right)$ of the action of the signals applied to the n-th compressor actuators. Under the influence of $u_{n,l}$ and $u_{n,r}$ sum, the characteristics of the air flow at the compressor outlet are formed. Their difference determines the intensity of the compressor operation. The current characteristics of the environment $\mathbf{x}_n = \left(x_{n,v},\ x_{n,\psi}\right)$, where $x_{n,v}$ and $x_{n,\psi}$ are current temperature and humidity of the air, respectively, are measured by a set of

sensors that form the output data vector $\mathbf{y}_n = \left(y_{n,v},\ y_{n,\psi}\right)$, where $y_{n,v}$ and $y_{n,\psi}$ results of measuring the current temperature and humidity. The controller converts $y_{n,v}$ and $y_{n,\psi}$ into digital codes, calculates the deviations of the values from those required for each n-th block and transfers them to the organizing controller, which generates regular local tasks for the compressor blocks.

The efficiency of swarm control (PPE components) is determined by the specifics of compressor control algorithms and the properties of von Neumann-type controllers as physical devices that generate time intervals between transactions when implementing algorithms. Due to the fact that the data processed by the controller is random, the control algorithm has a number of branch points, and the computational complexity of different branches can be different, the time intervals between transactions are also random. Thus, the PPE control system model can be represented as a $(N + 1)$-parallel semi-Markov process [15]

$$\tilde{\mu} = \{\mu_0(t),\ \mu_1(t),\ ...,\ \mu_n(t),\ ...,\ \mu_N(t)\},$$

where $\mu_0(t)$ describes the organizing controller operation; $\mu_n(t)$, $1 \le n \le N$, describes the operation of n-th controller; t is the time;

$$\mu_n = \{A_n,\ \mathbf{h}_n(t)\}.$$

In $A_n = \left\{a_{1(n)},\ ...,\ a_{j(n)},\ ...,\ a_{J(n)}\right\}$ are states of semi-Markov process $\mu_n(t)$, $\mathbf{h}_n(t) = \mathbf{g}_n(t) \otimes \mathbf{p}_n = \left[h_{j(n),l(n)}\right]$ is the $J(n) \times J(n)$ semi-Markov matrix; $\mathbf{g}_n(t) = \left[g_{j(n),l(n)}(t)\right]$ is the $J(n) \times J(n)$ matrix of pure time densities; $\mathbf{p}_n = \left[p_{j(n),l(n)}\right]$ is the $J(n) \times J(n)$ stochastic matrix.

In the context of the problem being solved, only the time intervals between transactions are significant, therefore $J(0) = 2N$, by organizing surveys of N sensors when entering data into the compressor control system and monitoring the quality of control over the characteristics of the air mixture; $J(n) = 4$, $1 \le n \le N$, according to the number of limbs.

The PPE control algorithm can have any structural complexity, but the principle of priority for obtaining data on the characteristics of the air environment immediately before the formation of control actions on the compressor components must be strictly observed. So, semi-Markov process may be reduced to process with the structure, shown on the Fig. 2a, where $a_{1(n)} \ldots a_{3(n)}$ are operators for reading data from sensors; $a_{4(n)} \ldots a_{6(n)}$ operators for writing data to the controller; DP – data processing operator of the formation of control actions in the compressor.

Transactions flow is shown on the Fig. 2b, where arrows with round tip indicate data inputs, while arrows with square tip indicate data outputs. Similar structure may be obtained from 0-th algorithm, describing organizing controller operation.

With use methods, described in [10–12, 22], for every neighboring pair of transactions time delays $\tau_{\xi(n),\zeta(n)}$, where $\xi(n), \xi(n) \in \{1(n),\ ...,\ 6(n)\}$, for $1 \le n \le N$, and $\xi(0), \xi(0) \in \{1(0),\ ...,\ N(0),\ N(0) + 1,\ ...,\ N(0) + N(0)\}$, determined, as well as the total operating time of the control program. Time intervals between transactions estimated, or either using «the three sigma rule» [23–25].

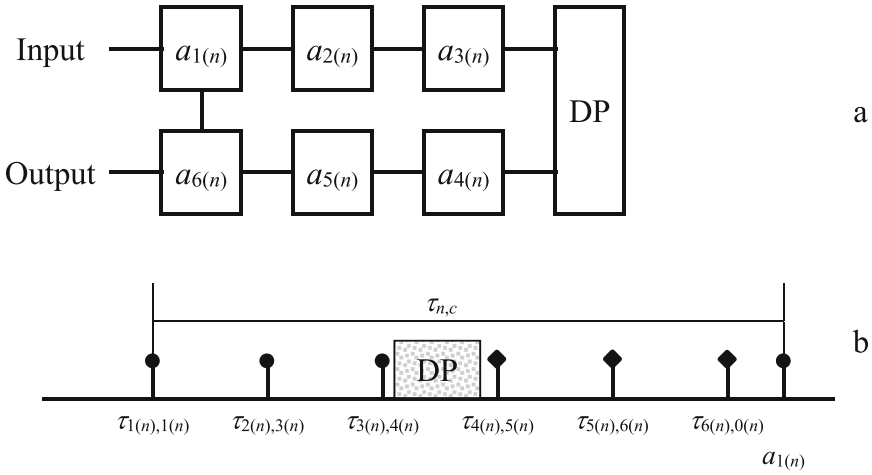


Fig. 2. Reduced structure of control algorithm

3 Results and Discussion

To confirm the correctness of the theoretical results, an analogous computational experiment was carried out, involving the control of the components of the PPE control system, consisting of five units [26–28]. The response of the compressor to the control action, which implies a change (increase or decrease) in air temperature by 1 °C and a simultaneous change (increase or decrease) in air humidity by 1% $\left[f_{n,v}(t),\ f_{n,\psi}(t)\right] = [\eta(t),\ \eta(t)]$ is shown on the Fig. 3.

The graphs presented in Fig. 3a show the change in air temperature ($T(t)$). The graphs presented in Fig. 3b show the change in air humidity ($H(t)$). Curves 1 are obtained when the delay in the formation of control actions in the compressor is equal to zero; curves 2 were obtained by estimating the delay without the «rule of three sigma» and equal to 0,02 s; curves 3 were obtained taking into account the «rule of three sigma» and is equal to 0,03 s. It can be seen that a decrease in the speed of the controller leads to overshoot and an increase in the time of the transient process.

On Fig. 4 shows the time characteristics of achieving the required characteristics when implementing a step-by-step PPE control algorithm. At each step, the controller sets the discreteness of the change in temperature and air humidity, which must be provided using the compressor control system for a given time.

Curve 4a shows the distribution density of the time of change in air temperature by 1 °C. Curve 4b shows the distribution density of the time of change in air humidity by 1%. Curve 4c shows the density distribution of the waiting time to achieve the required characteristics of the air environment. As can be seen from the graphs, the waiting time depends on the consistency of the implementation of control actions that take into account the values of temperature and humidity at all points where the air temperature and humidity sensors are located.

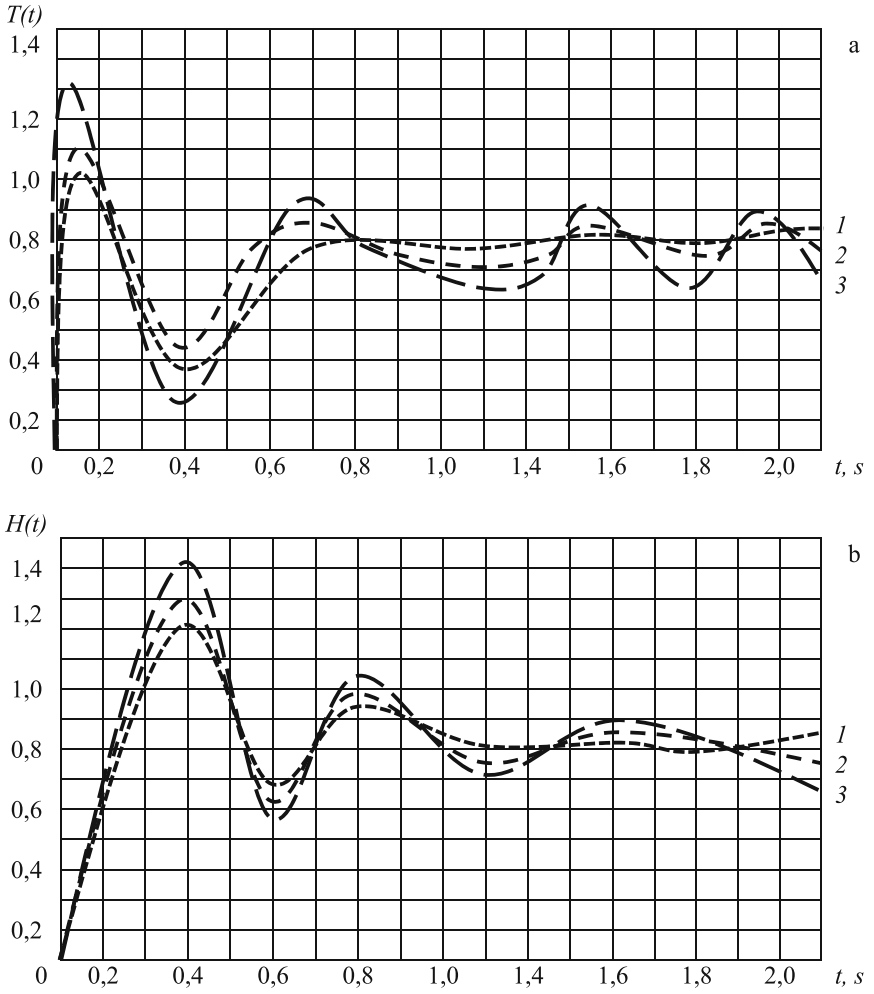


Fig. 3. Reaction of the swarm unit on the Heaviside function

Therefore, the control algorithm embedded in the controller must work in such a way as to increase the consistency of the control of compressor characteristics during the entire period of application of the PPE.

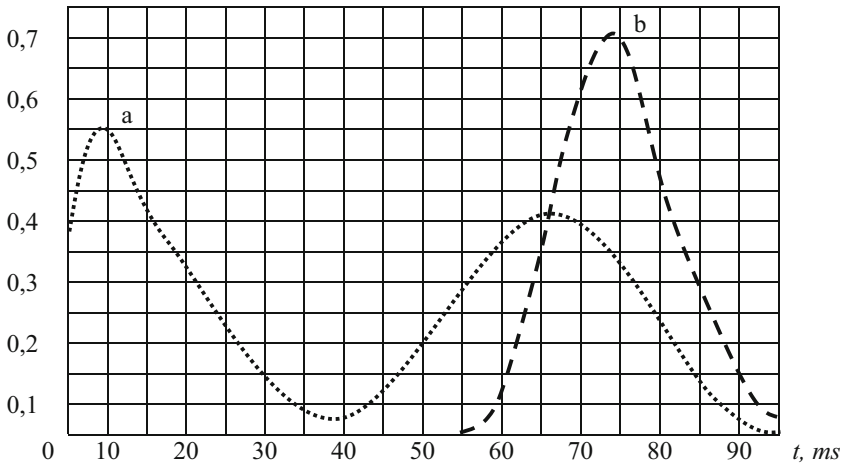


Fig. 4. Time distribution of getting the milestone by *n*-th unit (a), by swarm as a whole (b) and the backlog in getting the milestone by the swarm

4 Conclusion

The concept of controlling the functioning of intelligent PPE from infectious diseases is proposed in the form of a two-level hierarchical system for controlling PPE components, which ensures optimal characteristics of the air mixture in the space between the PPE and the body of the PPE user.

The implementation of control actions is based on the theory of semi-Markov processes, taking into account the time characteristics of von Neumann type controllers, which are the basis of the control system hardware.

Acknowledgements. The work was supported by the Russian Science Foundation, project no. 22-29-00808.

References

1. Batov, V.E.: Assessment of the functional state of military medical personnel when using personal protective equipment during the covid-19 pandemic. Med.-Biol. Soc.-Psychol. Prob. Saf. Emerg. Situat. **1**, 82–88 (2022)
2. Khristoforova, E.L., Nozdrevatykh, I.V.: The use of personal respiratory protection equipment in the context of the covid-19 pandemic. Tuberc. Soc. Signif. Dis. **1**(10), 58–59 (2022)
3. Bogomolov, A.: Information technologies of digital adaptive medicine. Inf. Autom. **20**(5), 1153–1181 (2021)
4. Larkin, E.V., Akimenko, T.A., Bogomolov, A.V.: Modeling the reliability of the onboard equipment of a mobile robot. Izvestiya of Saratov University. Mathematics. Mechanics. Informatics **21**(3), 390–399 (2021)
5. Larkin, E., Akimenko, T., Bogomolov, A., Krestovnikov, K.: Mathematical model for evaluating fault tolerance of on-board equipment of mobile robot. Smart Innov. Syst. Technol. **187**, 383–393 (2021)

6. Aström, J., Wittenmark, B.: Computer Controlled Systems: Theory and Design. Tsinghua University Press. Vashington, Prentice Hall (2002)
7. Landau, I.D., Zito, G.: Digital Control Systems, Design, Identification and Implementation. Springer, Berlin (2006). https://doi.org/10.1007/978-1-84628-056-6
8. Larkin, E., Privalov, A., Bogomolov, A., Akimenko, T.: Model of digital control system by complex multi-loop objects. AIP Conf. Proc. **2700**, 030009 (2023)
9. Bobtsov, A.A., Kapitonov, A.A., Nikolaev, N.A.: Control over the output of nonlinear systems with unaccounted-dynamics. Autom. Remote. Control. **71**(12), 2497–2504 (2010)
10. Larkin, E.V., Akimenko, T.A., Bogomolov, A.V.: Self-organizing mobile robots swarm movement control simulation. In: Tan, Y., Shi, Y., Niu, B. (eds.) ICSI 2022. LNCS, vol. 13345, pp. 56–65. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-09726-3_6
11. Larkin, E., Bogomolov, A., Privalov, A.: Digital control by robot manipulator with improved rigidity. Smart Innov. Syst. Technol. **329**, 45–59 (2023)
12. Larkin, E., Privalov, A., Bogomolov, A., Akimenko, T.: Digital control of continuous production with dry friction at actuators. Smart Innov. Syst. Technol. **232**, 427–436 (2022)
13. Janssen, J., Manca, R.: Applied Semi-Markov Processes. Springer, Berlin (2006). https://doi.org/10.1007/0-387-29548-8
14. Jiang, Q., Xi, H.-S., Yin, B.-Q.: Event-driven semi-Markov switching state-space control processes. IET Control Theory Appl. **12**(6), 1861–1869 (2012)
15. Girault, C., Valk, R.: Petri Net technology for Systems Engineering. A Guide to Modelling, Verification and Applications. Springe, Berlin (2001). https://doi.org/10.1007/978-3-662-05324-9
16. Bubeev, Yu.A., Vladimirskiy, B.M., Ushakov, I.B., Usov, V.M., Bogomolov, A.V.: Mathematical modelling of spread covid-19 epidemic for preventive measures to protect life and health of elderly. Bull. South Ural State Univ. Ser. Math. Model. Program. Comput. Softw. **14**(3), 92–98 (2021)
17. Soldatenko, S., Bogomolov, A., Ronzhin, A.: Mathematical modelling of climate change and variability in the context of outdoor ergonomics. Mathematics **9**(22), 2920 (2021)
18. Maistrou, A.I., Bogomolov, A.V.: Technology of automated medical diagnostics using fuzzy linguistic variables and consensus ranking methods. IFMBE Proc. **25**(7), 38–41 (2009)
19. Golosovskiy, M.S., Bogomolov, A.V., Terebov, D.S., Evtushenko, E.V.: Algorithm to adjust fuzzy inference system of Mamdani type. Bull. South Ural State Univ. Ser. Math. Mech. Phys. **10**(3), 19–29 (2018)
20. Tobin, D., Bogomolov, A., Golosovskiy, M.: Model of organization of software testing for cyber-physical systems. Stud. Syst. Decis. Control **418**, 51–60 (2022)
21. Golosovskiy, M., Bogomolov, A., Balandov, M.: Algorithm for configuring sugeno-type fuzzy inference systems based on the nearest neighbor method for use in cyber-physical systems. Stud. Syst. Decis. Control **417**, 83–97 (2022)
22. Kobayashi, H., Marl, B.L., Turin, W.: Probability, Random Processes and Statistical Analysis. Cambridge University Press, London (2012)
23. Pukelsheim, F.: The three sigma rule. Am. Stat. **48**(2), 88–91 (1994)
24. Yeh, Y.-C., Chu, Y., Chiou, C.W.: Improving the sampling resolution of periodic signals by using controlled sampling interval method. Comput. Electr. Eng. **40**(4), 1064–1071 (2014)
25. Pavlov, A.V.: About the equality of the transform of Laplace to the transform of Fourier. Issues of Analysis **5**(4), 21–30 (2016)
26. Larkin, E., Akimenko, T., Bogomolov, A., Sharov, V.: Reliability of robot's controller software. In: Ronzhin, A., Sadigov, A., Meshcheryakov, R. (eds.) ICR 2023. LNCS, vol. 14214, pp. 289–299. Springer, Cham. (2023). https://doi.org/10.1007/978-3-031-43111-1_26

27. Larkin, E.V., Akimenko, T.A., Bogomolov, A.V.: The swarm hierarchical control system. In: Tan, Y., Shi, Y., Luo, W. (eds.) ICSI 2023. LNCS, vol. 13968, pp. 30–39. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-36622-2_3
28. Larkin, E., Privalov, A., Bogomolov, A., Akimenko, T.: Model of digital control system by complex multi-loop objects. In: AIP Conference Proceedings, vol. 2700, p. 030009 (2023)

Software for Urban Space Researching with Public Data for Irkutsk

O. A. Nikolaychuk[1(✉)], D. E. Kosogorov[1,2], and Yu. V. Pestova[1]

[1] Matrosov Institute for System Dynamics and Control Theory of Siberian Branch of Russian Academy of Sciences, Irkutsk, Russia
nikoly67@mail.ru
[2] Irkutsk State University, Irkutsk, Russia

Abstract. The research of urban space can help in city management by identifying patterns in the development of the city and finding the strengths and weaknesses of individual places and entire areas. The purpose of this work is to create a software (tool) for research. Data for the tool is gathered from open Internet data sources using parsing methods. After the gathering data is processed, the values are normalized, the current indicators are calculated, and reviews are analyzed with sentiment analysis methods. The processing results are provided to the user in the form of interactive dashboards that combine visualizations necessary for work. This tool can be used not only by people involved in urban space management but also by tourists to get a clear picture of which places are more popular and which ones are the best to visit based on their queries.

Keywords: Attractors · Data Gathering · Data Normalization · Sentiment Analysis · Data Visualization

1 Introduction

The research on urban space is a comprehensive analysis of urban space with the aim of finding strategies for city development, which includes social, economic, infrastructure, and other areas. The results of these studies are targeted at specialists, such as those in urban management, as well as city residents interested in their lives. The research process may involve analyzing official statistics, survey results from residents, web resources, and social networks.

Researchers note a rapid growth of data containing geolocation information since any human activity leaves a digital trace (or digital footprint is a unique set of actions on the internet or on digital devices). This is facilitated by networks of stationary and mobile sensors, such as smartphones, GPS, credit cards, passes, information created in social networks, and other web resources. Such data with high spatial and temporal resolution allows one to study the dynamics of cities using constantly updated information and, in the future, will provide the possibility of analyzing cities as a whole [1].

One way to study urban space is to explore urban attractions based on the analysis of digital traces left by residents and tourists. Attractions (attractors) are locations that

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 352–362, 2024.
https://doi.org/10.1007/978-3-031-53552-9_31

attract people around them. For example, these can be hotels, gas stations, shops, museums, temples, parking lots, train stations, sources of fresh water, popular photo spots, monuments, etc. Attractions always have specific geographic coordinates (latitude and longitude) and may also contain additional information, such as name, category, address, business hours, website, etc.

Similar studies are currently relevant and being conducted in various countries, such as the study of East Java based on TripAdvisor reviews [2], New York using data from Telegram [3], cities in Italy based on data from social platforms like Flickr, Facebook, and Twitter [4], tourist centers in Spanish cities based on data from Foursquare, Google Places, Twitter, and Airbnb [5], in solving urban planning and development tasks, such as assessing the quality of life of urban populations, investigating the emotional perception of the urban environment, and implementing the conception of "smart city" [6].

The technology of creating interactive web resources is used to implement research tools where, based on accumulated data, users can obtain analysis results in the form of graphs and other visualization options (word clouds, heat maps, etc.). In urban studies, examples of such services include: Imprecity (https://imprecity.ru/analytics): provides an emotions map of St. Petersburg, where users can mark points on the map, leave comments, and indicate corresponding moods; ChronotopeCity (https://chronotopecity.habidatum.com/#introduction) allows one to study and track street load in Moscow, including specific categories like restaurant workload.

Based on geolocation data, various tasks can be solved [1], for example: identifying the most visited places in the city and making decisions on choosing accommodation for tourists, optimizing the placement of services or commercial activities, identifying the most visited places, which is necessary for marketing research, studying the density of reviews and photos to identify both unappealing tourist areas and directions of urban administration's activities for the development of these areas, identifying the most visited places in the city for vacation planning, reporting on the most popular streets and places, and consequently managing tourist flows, as well as determining and offering new tourist routes.

In Russia, a methodology for creating the Quality of Urban Environment Index has been developed to identify the competitive advantages of cities, as well as the constraints preventing their development, current problems, and prospective development directions, and to determine the level of quality of the urban environment of cities [7]. This methodology is used to calculate and visualize the indices of cities based on 36 indicators, which are publicly available on the website of the Index of Cities (https://индекс-городов.рф/). The Quality of Urban Environment Index assesses the condition of various aspects of the urban environment, including housing and surrounding spaces, road networks, green spaces, public amenities, and the overall urban environment. Some indicators are calculated based on data from social networks, such as the number of attraction centers for the population and the attractiveness of green areas (measured by the number of photos). Other indicators, such as the diversity of cultural and recreational infrastructure and services in residential and green areas, are assessed solely based on the quantity of services and infrastructure elements without considering their quality. This can only be determined through the analysis of ratings and consumer reviews. The main drawback of this service, from the perspective of urban environment

researchers, is the lack of detailed information about specific areas of cities, which would allow residents to assess the situation and participate in the development of their own city.

The methodology for estimating the quality index of the urban environment is being developed in separate cities of the Russian Federation, in particular using the digital footprint of the Yandex.Maps platform and the VKontakte social network [8].

The aim of this work is to develop a tool for studying urban space in order to analyze and visualize points of attraction based on data from open sources (public data). To achieve this proposal, the following tasks need to be solved:

- to determine a data model for the urban environment to identify points of attraction and methods of data collection,
- to generate datasets about the urban environment,
- to analyze the data obtained using methods such as data processing, quantitative and qualitative analysis, and text analysis to evaluate attractive objects in the urban environment through ranking,
- to develop data visualization panels to present and analyze information to users without specialized programming skills.

The city of Irkutsk has been chosen as the target city for studying the urban environment. The city is located in Eastern Siberia, on the banks of the Angara and Irkut rivers, 66 km away from Lake Baikal. The population of the city is approximately 600,000, making it the sixth largest city in Siberia. Inbound tourism to Irkutsk is mainly transit-oriented, as tourists pass through the city on their way to Lake Baikal. Foreign tourists typically stay in the city for an average of 1.5–2 days, with the majority of them being travelers from China, Germany, and Japan. 70% of the tourist flow to the city occurs during the summer season.

This work has been supported by the Russian Science Foundation, Project No. 23-28-00844 (Monitoring of regional tourism based on analysis of data from open sources).

2 Data Sources

The chosen data sources are OpenTripMap, Overpass Turbo, and Google Maps, which are mapping services. Based on the selected data sources and formulated goal, a general data information model (see Fig. 1) is created, which includes the following concepts: urban space object: object identifier, name, object categories, data source, URL address, physical address, coordinates, quantitative indicators (photos, reviews, rating); object visitation: object identifier, day, time, occupancy percentage; anonymous reviews about the object: review identifier, object and user IDs, review text, rating, publication date; photos from reviews: photo and review identifiers, URL: address of the photo.

Categories (types) of objects are defined based on the OpenTripMap category hierarchy: shops, foods, transport, amusements, sport, accommodations, religion, natural, cultural, historic, industrial facilities, and architecture.

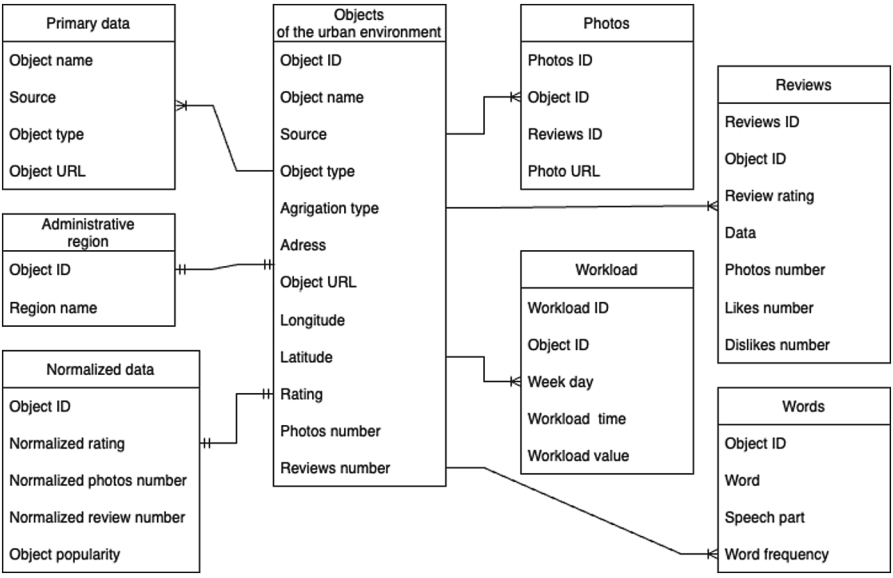


Fig. 1. A general data information model

3 Methods

3.1 Data Collection Methods

The following methods can be used to get data from web resources:

- online surveys,
- ready-made databases,
- API (Application Programming Interface),
- parsing (web scraping).

Online surveys differ from regular surveys in that the questionnaire is provided to the person through a website or messenger, which simplifies both the submission and filling out of the survey as well as obtaining the results. Obtaining data in this way takes a lot of time. For example, the service Imprecity uses this method to collect feedback, which is then used to create a mood map.

Ready-made databases allow for immediate access to a large volume of already structured data, but finding suitable databases freely available is quite difficult.

The Application Programming Interface (API) is an interface for exchanging data between applications. To obtain data using an API, the user must submit a request to the service and receive a structured response. This method is easy to implement, but services rarely provide the ability to obtain data through an API for free and without limitations. Collecting data through an API usually does not allow for obtaining data such as reviews and ratings.

Parsing is the automated collection and systematization of information from open sources using scripts when data is extracted directly from the markup of the website. This

method is the most difficult to implement, but it allows one to get a complete volume of data for effective analysis of the urban environment.

This work uses a combination of methods: the API and parsing (Fig. 2) for data collection. The API method of data sources Opentripmap.io and Overpass Turbo forms a database of places: information about the geographical location of an object, its name, address, category, and textual description.

The Google Maps data source is used to obtain user metrics and characteristics of places, including traffic (load percentage), ratings, reviews, photos, and their quantitative values. This web resource is a dynamic data source, meaning that all data is loaded into the web resource during the user's interaction with the browser. Therefore, the Selenium library (Python) is used for data collection automation through web scraping. It creates an agent that simulates user actions on the page and extracts information from the DOM structure.

After collecting the data, data preprocessing is conducted: collection errors are corrected, duplicate places are removed, and identified points are linked to administrative districts of the city (see Fig. 2).

3.2 Data Analysis Methods

Data analysis is proposed to be carried out based on methods of analyzing quantitative indicators and texts.

Quantitative indicators are normalized on the interval [0, 1] before use based on the formula for linear normalization of data:

$$f(X) = (X - \min(X))/(\max(X) - \min(X)), \tag{1}$$

where X is the value of the indicator.

This approach is common and can be implemented using ready-made functions, such as the MinMax() function in the scikit-learn library for Python [10].

To identify attraction points and their characteristics, indicators are proposed that allow one to assess the popularity and comfort of visiting these points.

In the methodology for forming the urban environment quality index, the calculation of the "number of attraction centers for the population" indicator is based on data on the number of photographs:

$$\mathrm{N_{str}}(F)/\mathrm{N_{str}}$$

where $N_{str}(F)$ is the number of streets most frequently mentioned when publishing photos (the total number of publications using photos taken on these streets accounts for 75 percent of all published photos taken within the city area), N_{str} is the total number of streets in the city.

The work suggests using indicators such as point rating, number of reviews, and number of photos as a way of evaluating attraction points, allowing us to get a more comprehensive assessment. Based on these indicators, a composite indicator is formed to characterize the attractiveness or popularity of the point by summing up weighted indicators [11]. In the scope of this study, the overall assessment is calculated as the sum

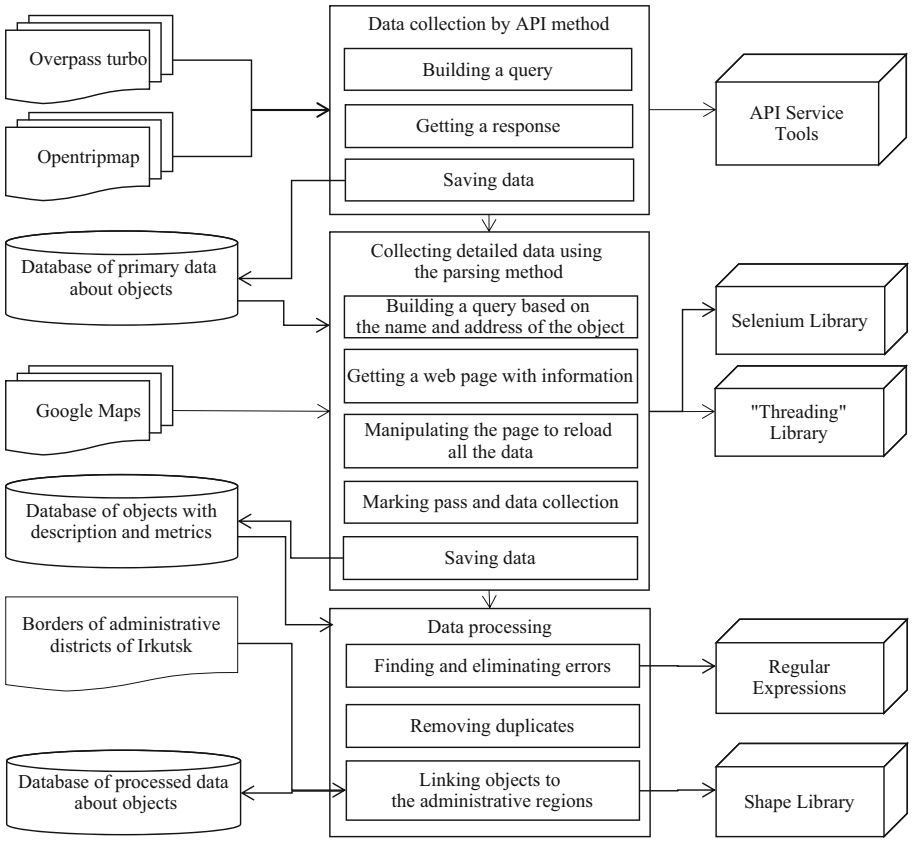


Fig. 2. Stages of data collection and processing

of the normalized rating, number of reviews, and number of photos, multiplied by their respective weights:

$$P_i = W_{rat}X_{rat} + W_{rev}X_{rev} + W_{pht}X_{pht}, \tag{2}$$

where X is the normalized criteria score, W is the weight or significance of the criteria (ranging from 0 to 3), *rat* is the rating, *rev* is the number of reviews, and *pht* is the number of photographs.

If the significance criteria are equal to 1, the popularity rating is measured on a scale from 0 to 3. Places with a value below 1 are not popular and are not considered attraction points (2785 places), between 1 and 1.7 (795 places) are considered popular or have the potential to become attraction points, and places with a value above 1.7 (15 places) are considered very popular and attraction points. The threshold values for classifying popularity were selected based on statistical processing.

The default weights for criteria are 1, but they can be changed by the user. For example, if a user considers the photogenicity of a place important, they can increase

the weight of this indicator compared to others or set it to zero. Therefore, the range of criteria values is chosen up to 3 so that it is possible to assign priority to each indicator.

The second indicator for evaluating points of attraction is the "comfort" of visiting a place. Within the scope of this work, visiting popular places when they are less crowded is considered comfortable. Therefore, places with high popularity and low crowding have the highest comfort rating during visiting in a specific period of time:

$$O_{i,d,t} = P_i / V_{i,d,t}, \tag{3}$$

where $Q_{i,d,t}$ is the comfort of i-place on day of the week d and time t, P_i is the popularity of i-place, $V_{i,d,t}$ is the occupancy of i-place on day of the week d and time t ([0, 100], %), time t is analyzed discretely with an interval of one hour.

This indicator shows the gradation of a comfortable stay by location and by time. Among places with available capacity, the indicator is measured at 0.2. Users are provided with a list of locations sorted by comfort in descending order, which means that the places with higher popularity but lower average occupancy are displayed at the top of the list. The user's selection of place categories, days, and time recalculates the indicator according to the specified filtration.

The next method of analyzing attractor points was the method of analyzing reviews [9]. The text of the reviews was divided into words, for which the normal form and part of speech were determined using the PyMorphy2 library (see Fig. 3). The obtained dataset of words in the normal form was also labeled with identifiers of locations, reviews, and review ratings. The numerical rating of a place (1–5 stars) from a review allows one

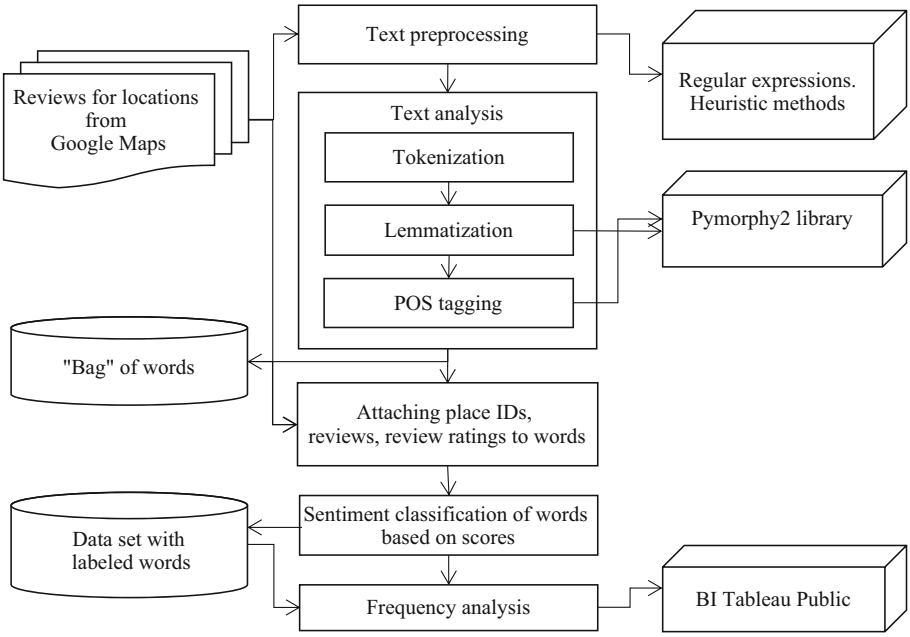


Fig. 3. Stages of text review processing and analysis

to indirectly determine the sentiment of the text, considering ratings of 1 and 2 stars as negative, 3 as neutral, and 4 and 5 as positive reviews. Then, a frequency analysis of words was conducted, which allowed us to get a characteristic description of the locations and some representation of them, with the review word rating being calculated as an average and reflecting the sentiment.

3.3 Data Visualization

To present the results of data analysis, it is necessary to use visualization methods. Possible visualization options include graphs, bar charts, tables, heat maps, word clouds, geographic maps, and pie charts. As part of this study, visualization panels have been

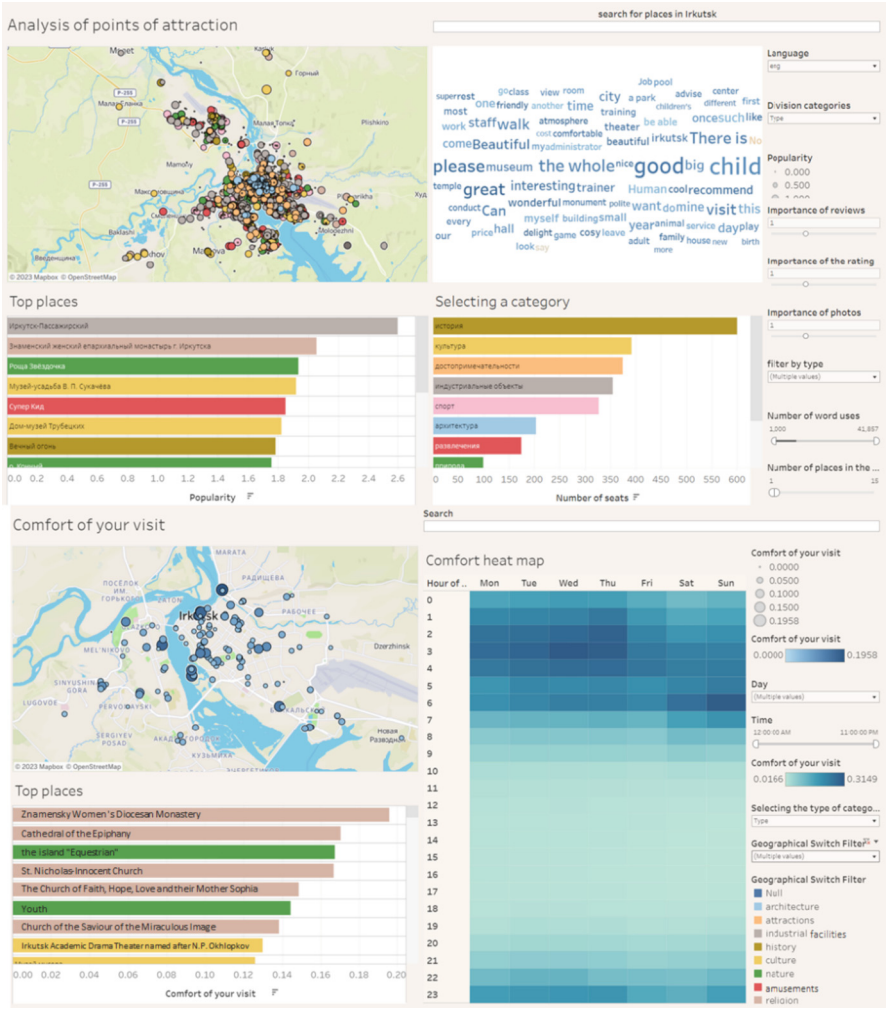


Fig. 4. The example of Information Analysis Results

developed. Visualization panels (or dashboards) are a thematic combination of several interactive diagrams, which allows for their management through interaction with all elements of the panel for filtering, sorting, and selecting parameters [12]. In this work, dashboards have been developed using the business intelligence software Tableau Public.

Visualization panels are designed to solve the following tasks (see Fig. 4):

- Analysis of the popularity of attractions:
- Analysis of the spatial location of popular attractions on the city map,
- Analysis of attraction categories based on the number of popular attractions (bar chart),
- Analysis of attraction popularity ratings with the ability to set weights for indicators used in the calculation (bar chart),
- Analysis of attractions based on word clouds extracted from reviews about these attractions.
- Analysis of the comfortability of visiting attractions:
- Analysis of the spatial location of comfortable attractions on the city map.
- Analysis of the comfortability ratings of attractions (bar chart),
- Analysis of the comfortability of attractions in terms of weekdays and hours of the day (heatmap).

4 Analysis Results

As a result of the research, data on the points of attraction in the city environment was obtained. The following most popular points have been identified: transportation - Irkutsk Passenger Railway Station; dining establishments: Pho ME, Rio Grande, and Kochevnik restaurant; entertainment centers: Yubileyny Shopping Center, Jam Mall, and Yarkomall Shopping Center; religious sites: Znamensky Monastery; museums: V.P. Sukachev Manor and Trubetskoy House-Museum; nature (parks): Zvezdochka Grove, Konny Island; monuments (memorial complexes): eternal flame and Babr Monument.

The tool allows one to analyze the points of attraction for each district of the city. For example, the central districts, being the historical center of the city, have the most popular points of attraction among landmarks: monuments (memorial complexes) dedicated to the Second World War, Yakov Pokhabov (the founder of the city), and Leonid Gaidai (a famous Russian film director born in the city); the religious site Znamensky Women's Diocesan Monastery, built a few decades after the foundation of Irkutsk Ostrog in 1689; the historical site Trubetskoy House-Museum, Decembrists of 1825; and the industrial site Art-Zavod Dorenberg, organizing space for creativity and creative entrepreneurship.

Based on the density maps, the density of attraction points for streets, quarters, and entire districts is estimated: Kirovsky District - Karla Marksa Street, Lenina Street, the district around Kirov Square, which forms the basis of the city's historical center, Oktyabrsky District - the 130th quarter, which consists of restored buildings that were predominant in the city in the 18th and 19th centuries, Sverdlovsky District - Lermontova Street, which runs along the Angara River and crosses a large part of the district, Kuibyshevsky District - Dekabrskikh Sobytii Street and the quarter around Volkonsky Square, Leninsky District - Akademika Obraztsova Street.

A word cloud allows you to obtain a characteristic of the object or category of objects being discussed based on reviews, as well as search for objects based on a selected characteristic. For example, by selecting the word “animal” in the object category “nature,” you can determine in which parks of the city they can be seen (see Fig. 5).

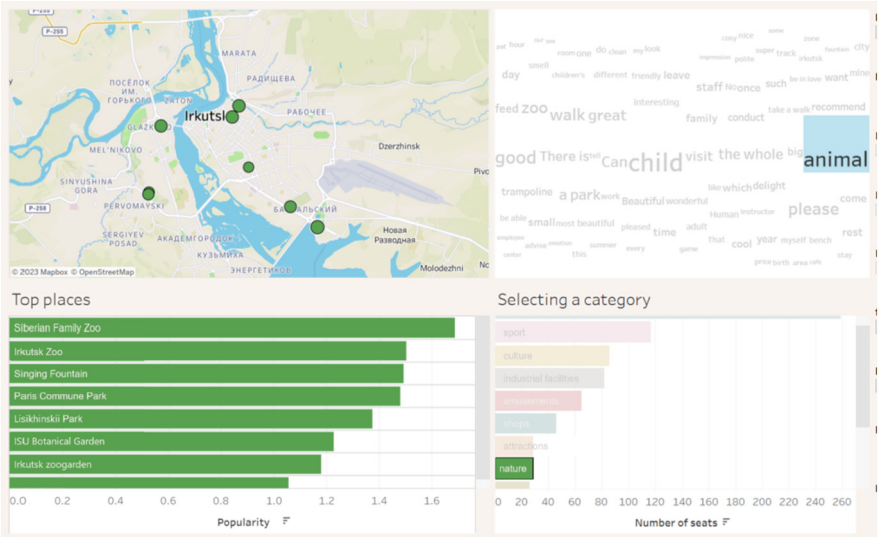


Fig. 5. Data search results for the “animal” keyword

5 Conclusions

The results of the work include implemented data collection algorithms for each data source (OpentripMap, Overpass Turbo, and Google Maps); created datasets of preprocessed lists of locations, their metadata (description and quantitative metrics, visitation, reviews, and photos); models of indicators for popularity and comfort of visiting points (places); a dataset of frequency analysis of review texts with linkage to user ratings; and interactive visualization panels for the analysis results.

As a result of the data collection, a dataset of points was formed consisting of 3595 records. Most of the places are in the Kirovsky district, the historical center of the city. Among the top 15 places in terms of “popularity,” there are no points from the Leninsky district, which is a “residential” district, or outside the city, which can be considered potential territories for the formation of new points of attraction. In terms of categories, historical and cultural places are more represented, which characterizes Irkutsk as a city with a cultural heritage.

For a comfortable stay, religious sites (Znamensky women’s diocesan monastery, Nikolai-Innokentievsky temple, and others) are recommended based on the analysis, as are natural sites (Konny Island and Youth) in the morning and evening.

The developed tool provides obtaining more complete and reliable data about the state of the urban environment, which in turn allows one to increase the efficiency of urban management and the level of orientation of the residents in the city's life. The collected data will be supplemented and used in further research of the urban space of Irkutsk city.

References

1. García-Palomares, J.C., Gutiérrez, J., Mínguez, C.: Identification of tourist hot spots based on social networks: a comparative analysis of European metropolises using photo-sharing services and GIS. Appl. Geogr. **63**, 408–417 (2015)
2. Satria, D., Wibowo, J.M.: Big Data Analysis of Sustainable Tourism Competitiveness in East Java Province. https://www.hippocampus.si/ISSN/2335-4194/14.189-203.pdf. Accessed 09 May 2023
3. Hu, F., Li, Z., Yang, C., Jiang, Y.: A graph-based approach to detecting the tourist movement patterns using social media data. https://www.researchgate.net/publication/325331868_A_graph-based_approach_to_detecting_the_tourist_movement_patterns_using_social_media_data. Accessed 17 Mar 2023
4. Giglio, S., Bertacchini, F., Bilotta, E., Pantano, P.: Using social media to identify tourism attractiveness in six Italian cities. Tour. Manage. **72**, 306–312 (2019)
5. Martí, P., García-Mayor, C., Serrano-Estrada, L.: Taking the urban tourist activity pulse through digital footprints. Curr. Issue Tour. **2**(24), 157–176 (2021)
6. Nenko, A.E., Kurilova, M.A., Podkorytova, M.I.: Analysis of emotional perception of territories and development of "smart city." Int. J. Open Inf. Technol. **11**(8), 128–136 (2020)
7. Decree of the Government of the Russian Federation dated March 23, 2019 No. 510-r (as amended on December 30, 2020) "On the Approval of the Methodology for the Formation of the Urban Environment Quality Index"
8. Kurcheeva, G.I., Babaykina, E.Yu.: Development of a methodology for applying digital footprint to monitoring indicators of population's quality of life. In: Babkin, A.V. (eds.) Digital Economy and Industry 5.0: Development in a New Reality, POLYTECH-PRESS Publishing House, St. Petersburg, pp. 381–407 (2022)
9. Medhat, W., Hassan, A., Korashy, H.: Sentiment analysis algorithms and applications: a survey. https://www.sciencedirect.com/science/article/pii/S2090447914000550. Accessed 14 Apr 2023
10. Patro, S.G.K., Sahu, K.K.: Normalization: A Preprocessing Stage. Electronic archive. Cornell University. arXiv:1503.06462. Accessed 15 May 2019
11. Klyushnikova, E.V., Shitova, E.M.: Methodological approaches to calculating the integral indicator, ranking methods. InnoCenter **1**(10), 4–18 (2016)
12. Pestova, Yu.V.: Data analysis from open sources for monitoring the tourist infrastructure objects of the Baikal region. In: Communication Technologies: Socio-Economic and Informational Aspects: Proceedings of the All-Russian Youth Scientific and Practical Conference, pp. 52–56. LLC CenterNauchService, Irkutsk (2022)

Core-Intermediate-Peripheral Index: Factor Analysis of Neighborhood and Shortest Paths-Based Centrality Metrics

Natarajan Meghanathan(✉)

Department of Electrical and Computer Engineering and Computer Science, Jackson State University, 1400 Lynch Street, Jackson, MS 39217, USA
natarajan.meghanathan@jsums.edu

Abstract. We perform factor analysis on the raw data of the four major neighborhood and shortest paths-based centrality metrics (Degree, Eigenvector, Betweenness and Closeness) and propose a novel quantitative measure called the Core-Intermediate-Peripheral (CIP) Index to capture the extent with which a node could play the role of a core node (nodes at the center of a network with larger values for any centrality metric) vis-a-vis a peripheral node (nodes that exist at the periphery of a network with lower values for any centrality metric). We conduct factor analysis (varimax-based rotation of the Eigenvectors) on the transpose matrix of the raw centrality metrics dataset, with the node ids as features, under the hypothesis that there are two factors (core and peripheral) that drive the values incurred by the nodes with respect to the centrality metrics. We test our approach on a diverse suite of 12 complex real-world networks.

Keywords: Factor analysis · Core nodes · Intermediate Nodes · Peripheral nodes · Centrality metrics · Varimax rotation

1 Introduction

The topological importance of nodes in complex networks has been analyzed in the literature from the perspectives of core-periphery structure and centrality metrics. While the core-periphery structure analysis of a network is more of a qualitative approach at a mesoscopic level, centrality metrics are designed to quantify the topological importance of individual nodes in a network. The core-periphery analysis of a network is aimed at categorizing a node as either a core node or a peripheral node. The current status quo in the literature on the definitions of core nodes and peripheral nodes is that the core nodes need to be of larger degree and form a highly dense backbone to which the low degree peripheral nodes are connected to; the peripheral nodes are expected to be not connected to other peripheral nodes as well. Some of the works (e.g., [1–3]) in the literature have suggested that high degree nodes need not always be core nodes; but they still analyze the core-periphery structure and quantify the extent of coreness of a node within the realms of the above model.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 363–372, 2024.
https://doi.org/10.1007/978-3-031-53552-9_32

Several of the existing algorithms to analyze the core-periphery structure of a network would fail to identify node 2 (in Fig. 1) of our running example graph of Sect. 2 as a core node and on the other hand, would overrate and classify nodes 3 and 6 as core nodes. A close look at the example graph of Fig. 1 would indicate that nodes 1 and 2 bring the two otherwise disconnected components of nodes ([1, 3, 5, 6, 9] and [2, 4, 7, 8, 10]) together and must be highly ranked as core nodes, whereas nodes 3 and 6 are to be treated as peripheral nodes, because their presence do not add any value to the network, other than increasing the degrees of nodes 1 and 5. We do not claim a core node has to be an articulation point [4] (whose removal would disconnect an otherwise connected graph); but, we do expect a core node to add value to the connectivity of the nodes in the network to justify the notion that the core nodes are to form the center and the peripheral nodes are at the boundary of the network.

In this paper, we seek to enhance the core-periphery model by proposing a more comprehensive definition for the core nodes and peripheral nodes on the basis of the values incurred by the nodes for both neighborhood-based and shortest paths-based centrality metrics. The four commonly studied centrality metrics spanning these two categories are: Degree (DEG) [5], Eigenvector (EVC) [6], Betweenness (BWC) [7] and Closeness (CLC) [8]. While DEG (measure of the number of neighbors of a node) and EVC (measure of the degree of a node as well as the degrees of its neighbors) are neighborhood-based metrics, BWC (measure of the extent with which the shortest paths between any two nodes in the network go through the node) and CLC (measure of the lengths of the shortest paths from the node to the rest of the nodes in the network) are shortest paths-based metrics.

Our premise is: nodes that are part of the shortest paths for several node pairs and located closer to the majority (if not all) of the nodes in the network are more likely to incur reasonably larger values for the neighborhood-based centrality metrics as well. Accordingly, we propose the following definitions for a core node and a peripheral node: a core node should form the center of the network by being closer to the rest of the nodes and located on the shortest paths between several node pairs; whereas the peripheral nodes form the boundary of the network by being far away from a majority of the nodes and barely located on the shortest paths for any node pairs. To the best of our knowledge, we have not come across any work that analyzes the core-periphery structure of a network by taking into consideration the values incurred by the nodes with respect to a comprehensive set of centrality metrics spanning both the neighborhood-based and shortest paths-based categories. Our work in this paper takes the latter direction.

Our hypothesis for this research is that the values incurred by a node with respect to different centrality metrics are majorly influenced by their location: either at the core or the periphery of a network. Our hypothesis stems from the observation that the core nodes incur significantly larger values for the centrality metrics (especially, BWC, CLC and DEG), while the peripheral nodes incur significantly lower centrality values. Though there are quite a few works in the literature (e.g., [1]) that quantify the extent of coreness of a node, none of these works take into consideration the shortest paths aspect and are heavily-based on degree centrality. This forms the motivation for our research. We propose to conduct factor analysis [9] on the centrality dataset (with respect to values incurred for the DEG, EVC, BWC and CLC metrics) of the nodes in a network and seek

to quantify the latent (hidden) factors (core or peripheral) that drive the values incurred by a node with respect to the different centrality metrics.

The rest of the paper is organized as follows: Sect. 2 explains the proposed centrality dataset-based factor analysis approach for core-periphery analysis along with a running example graph. Section 3 presents the results of running the proposed factor analysis approach on a suite of 12 diverse complex real-world networks and classifies them as heavy with respect to one or two of the three classes of nodes (core, peripheral and intermediate) that we identify from the factor analysis approach. Section 4 discusses related work in the literature and highlights the uniqueness of our work. Section 5 concludes the paper and presents plans for future work.

2 Factor Analysis of Centrality Dataset

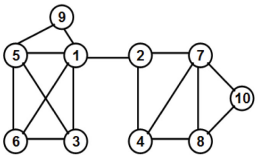
Factor analysis [9] is a widely used approach in machine learning to quantify the hidden factors that are behind the values incurred for the different features (columns) in a dataset (matrix) of records. For the problem in hand, the features are the nodes and we hence conduct factor analysis on the transpose matrix (that will have four rows, each pertaining to a centrality metric: DEG, EVC, BWC and CLC and n columns, where n is the number of nodes in the network) of the centrality dataset. Figure 1 presents a toy 10-node graph (and the transpose matrix of its centrality dataset) that we will use as a running example in this section to illustrate our proposed procedure.

We first obtain the covariance matrix (see Fig. 1) of the transpose matrix of the centrality dataset and determine its Eigenvalues and Eigenvectors [10]. Since we hypothesize that there are only two factors behind the centrality values incurred by the nodes, we retain only the Eigenvectors (EVs; see Fig. 2) corresponding to the largest and second largest Eigenvalues. The entries in these two Eigenvectors (referred to respectively as the first principal and second principal Eigenvectors) are considered as the initial loadings for the nodes (features). We build a two-dimensional coordinate system (see Fig. 2) of the Eigenvectors with these node loadings as the data points.

We seek to synchronously rotate the two Eigenvector coordinate-axes (that are orthogonal to each other) in such a way that the maximum of the data points are aligned on either of them (a procedure referred to as Varimax rotation [10] in the literature). Varimax rotation aims to maximize the communality score (sum of the squares of the loadings) for each node by going through repeated orthogonal rotations of the Eigenvector coordinate-axes. We conducted Varimax rotation using the relevant libraries available in Python (Pandas) [11]. The axes in the resulting rotated coordinate system (see Fig. 2) correspond to the two factors (core and peripheral); more specifically, we treat the axis where nodes with larger values for BWC align with as the vertical Y-axis (referred to as the core-axis) and the axis where nodes with lower or zero BWC values align with as the horizontal X-axis (referred to as the peripheral-axis). The coordinates (final loadings; see Fig. 2) of the nodes in such a rotated coordinate system are expected to be either close to (1, 0): if the node is a peripheral node or (0, 1): if the node is a core node.

Core-Intermediate-Peripheral (CIP) Index: With the core-axis being the vertical Y-axis (and the peripheral-axis being the horizontal X-axis), the Core Intermediate Peripheral: CIP Index value for a node could be construed as a quantitative estimate of the extent

with which a node could serve as a core node in the network. We determine the CIP Index as the tan (tangent) angle of the line joining the origin (0, 0) and the coordinates for a node in the peripheral axis-core axis coordinate system. The angle is measured anti-clockwise (clockwise) from the peripheral-axis for those coordinates in the first/second quadrants (fourth quadrant). If the CIP index for a node is closer to 90° (or closer to 0°), the node could be considered a core node (or a peripheral node). Figure 3 illustrates the CIP angle measurements for nodes 2 (a core node), 4 (an intermediate node) and 9 (a peripheral node) in the example graph.



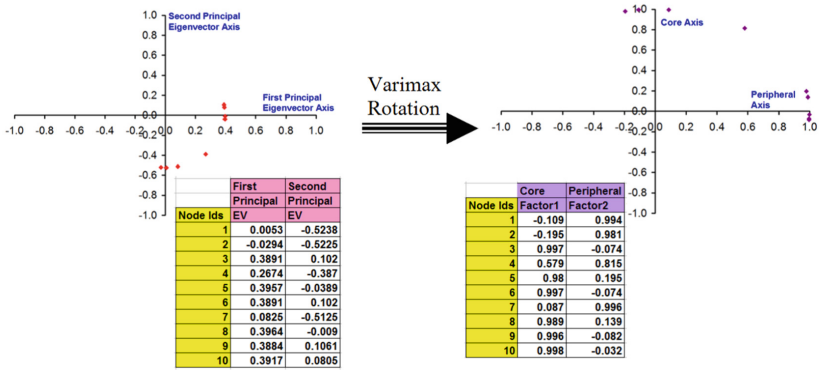
Graph

	1	2	3	4	5	6	7	8	9	10
DEG	5	3	3	3	4	3	4	3	2	2
EVC	0.5127	0.2443	0.3949	0.1564	0.4579	0.3949	0.1756	0.1208	0.2807	0.0857
BWC	21	20	0	3	1	0	9.5	0.5	0	0
CLC	0.0667	0.0667	0.0476	0.0526	0.0500	0.0476	0.0556	0.0417	0.0455	0.0400

Transpose Matrix of the Centrality Dataset

	1	2	3	4	5	6	7	8	9	10
1	1	0.9962	-0.181	0.7471	0.0879	-0.181	0.9806	0.0297	-0.189	-0.141
2	0.9962	1	-0.266	0.6863	0.0003	-0.266	0.9599	-0.058	-0.274	-0.227
3	-0.181	-0.266	1	0.5161	0.9637	1	0.0133	0.9744	1	0.9959
4	0.7471	0.6863	0.5161	1	0.726	0.5161	0.8627	0.6866	0.5091	0.5526
5	0.0879	0.0003	0.9637	0.726	1	0.9637	0.2798	0.9959	0.9615	0.9713
6	-0.181	-0.266	1	0.5161	0.9637	1	0.0133	0.9744	1	0.9959
7	0.9806	0.9599	0.0133	0.8627	0.2798	0.0133	1	0.2246	0.0052	0.0552
8	0.0297	-0.058	0.9744	0.6866	0.9959	0.9744	0.2246	1	0.9725	0.9853
9	-0.189	-0.274	1	0.5091	0.9615	1	0.0052	0.9725	1	0.9954
10	-0.141	-0.227	0.9959	0.5526	0.9713	0.9959	0.0552	0.9853	0.9954	1

Fig. 1. Visualization of the Running Example Graph; Transpose Matrix of its Centrality Dataset and its Covariance Matrix



Node Ids	First Principal EV	Second Principal EV
1	0.0053	-0.5238
2	-0.0294	-0.5225
3	0.3891	0.102
4	0.2674	-0.387
5	0.3957	-0.0389
6	0.3891	0.102
7	0.0825	-0.5125
8	0.3964	-0.009
9	0.3884	0.1061
10	0.3917	0.0805

Node Ids	Core Factor1	Peripheral Factor2
1	-0.109	0.994
2	-0.195	0.981
3	0.997	-0.074
4	0.579	0.815
5	0.98	0.195
6	0.997	-0.074
7	0.087	0.996
8	0.989	0.139
9	0.996	-0.082
10	0.998	-0.032

Fig. 2. Varimax Rotation of the Eigenvector-Axes to the Factor-Axes

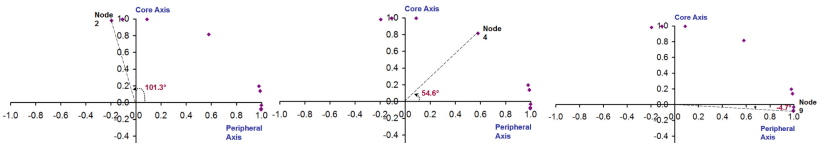


Fig. 3. Sample Core-Intermediate-Peripheral (CIP) Index Angle Measurements for Nodes in the Example Graph

3-Element CIP Bins-Fraction Tuple: We also propose a binning approach (with interval of 10 degrees: <0, 0..10, 10…20, etc.., 80…90 and ≥ 90) to group the nodes based on their CIP values. We treat the bins for CIP values falling in either of these two ranges [80,…, 90) and [90,….) as those corresponding to the core nodes and the bins for CIP values falling in either of these two ranges […, 0) and [0,…, 10) to correspond to the peripheral nodes. We categorize the rest of the CIP bins to belong to nodes that are neither core nor peripheral (we refer to such nodes as intermediate nodes). We observe the centrality values for such intermediate nodes to be neither too high nor too low. For a given network, we determine a 3-element CIP bins-fraction tuple that would comprise the fractions of the nodes in the core, peripheral and intermediate classes. If the value for any of these three fractions is 0.5 or more, we categorize the network as heavy with respect to that class. If all the three fractions in the 3-element CIP bins-fraction tuple for a network are less than 0.5, we categorize the network as heavy with respect to the classes with the top two fraction values.

	Core	Peripheral	CIP Index	Ranking of Nodes
Node Ids	Factor1	Factor2	Tan Angle	Extent of Coreness
1	-0.109	0.994	96.3	2
2	-0.195	0.981	101.3	1 (Top most rank)
3	0.997	-0.074	-4.2	8
4	0.579	0.815	54.6	4
5	0.98	0.195	11.2	5
6	0.997	-0.074	-4.2	8
7	0.087	0.996	85.0	3
8	0.989	0.139	8.0	6
9	0.996	-0.082	-4.7	10 (Bottommost rank)
10	0.998	-0.032	-1.8	7

Bins	Count	
< 0	4	5 Peripheral Nodes [3, 6, 8, 9, 10]
≥ 0… < 10	1	
≥ 10… < 20	1	2 Intermediate Nodes [4, 5]
≥ 20… < 30	0	
≥ 30… < 40	0	
≥ 40… < 50	0	
≥ 50… < 60	1	
≥ 60… < 70	0	
≥ 70… < 80	0	
≥ 80… < 90	1	3 Core Nodes [1, 2, 7]
≥ 90	2	

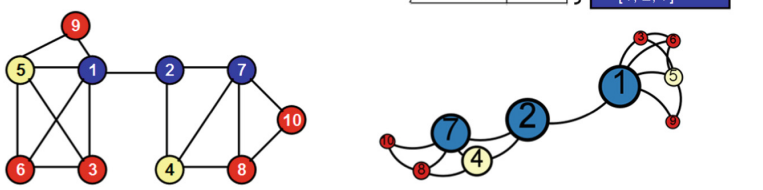


Fig. 4. Ranking, Classification and Visualization of the Nodes as Core, Intermediate and Peripheral Nodes based on their CIP Index Values

Figure 4 presents the CIP index values for all the nodes in the network as well as ranks them on the basis of the extent of coreness. Figure 4 also presents a count of the number of nodes falling in each of the CIP index range bins: we observe 3 of the 10 nodes to be core, 2 nodes to be intermediate and 5 nodes to be peripheral. The 3-element CIP bins-fraction tuple for the example graph is thus [3/10, 2/10, 5/10] = [0.3, 0.2, 0.5]. Since the fraction of peripheral nodes satisfies the criteria of being 0.5 or above, we conclude the example graph to be peripheral-heavy. In addition, Fig. 4 also visually illustrates the core-peripheral structure of the graph in a typical layout as well as using the Yifan proportional layout algorithm [14] (in both the layouts, the node color corresponds to the class of the node). In the Yifan Hu proportional layout, the node size is proportional to their CIP index values as well.

3 Evaluation with Real-World Networks

We present the results obtained by running the proposed factor analysis procedure of Sect. 2 on a suite of 12 diverse real-world networks, ranging from pure random networks [12] to extremely scale-free networks [13]. The diversity among these networks is captured in the form of the spectral radius ratio for node degree (represented as λ_{sp}) [15], a metric that seamlessly captures the variation in node degree independent of the number of nodes and edges in the network. The λ_{sp} value [15] for a network is the ratio of the largest Eigenvalue of the 0–1 adjacency matrix of the network and the average node degree. Table 1 presents the name, number of nodes and edges and the λ_{sp} value of the networks as well as the results (the 3-element bins-fraction tuple and the network classification) for the 12 real-world networks. Most of these real-world networks have been used as benchmark networks in several studies related to complex network analysis. For example, the well-studied Karate Network (Net 6) incurs a 3-element bins fraction tuple of [0.35, 0.21, 0.44] and accordingly the network is classified as peripheral/core-heavy: implying that the largest fraction of nodes in the network are peripheral nodes and the next largest fraction of nodes are the core nodes.

We observe the random networks (US Football Network and the Taro Exchange Network) to be intermediate-heavy. The rest of the real-world networks are scale-free networks with different levels of variation in node degree. We observe the scale-free networks with relatively lower λ_{sp} values to be predominantly core-heavy. Among the five real-world networks, with λ_{sp} values ranging from 1.21 to 1.73, we observe three of them to be core-heavy, one network to be intermediate/core-heavy and the other network to be peripheral/core-heavy. On the other hand, all the five scale-free real-world networks with a relatively larger variation in node degree (i.e., $\lambda_{sp} \geq 1.82$) are observed to be peripheral-heavy.

Figure 5 presents a visualization of each of the 12 real-world networks (identified with the Net #s used in Table 1), in the increasing order of their λ_{sp} values. All the real-world networks are displayed per the Yifan Hu proportional layout algorithm run in Gephi [16]. We also follow the same coloring convention for the core, intermediate and peripheral nodes as is shown in Fig. 4 (i.e., the core, peripheral and intermediate nodes are colored in blue, red and whitish-yellow respectively). We observe the real-world networks to showcase the following trend: random networks are intermediate-heavy;

Table 1. Real-World Networks used for Centrality-based Factor Analysis and their 3-Element Bins CIP Fractions Tuple-based Classification

Net-ID	Network Name	# Nodes	# Edges	λsp	[C, I, P] Fractions	Network Class
Net-1	US Football Net	115	613	1.01	[0.00, 1.00, 0.00]	Intermediate-heavy
Net-2	Taro Exchange Net	22	39	1.06	[0.36, 0.59, 0.05]	Intermediate-heavy
Net-3	Flying Teams Net	48	170	1.21	[0.50, 0.35, 0.15]	Core-heavy
Net-4	Dolphin Net	62	159	1.40	[0.68, 0.16, 0.16]	Core-heavy
Net-5	Band Jazz Net	198	2742	1.44	[0.33, 0.46, 0.21]	Intermed/Core-heavy
Net-6	Karate Net	34	78	1.47	[0.35, 0.21, 0.44]	Peripheral/Core-heavy
Net-7	Adjacency Noun Net	112	425	1.73	[0.62, 0.26, 0.12]	Core-heavy
Net-8	Les Miserables Net	77	254	1.82	[0.28, 0.14, 0.58]	Peripheral-heavy
Net-9	Copper Field Net	87	406	1.83	[0.11, 0.39, 0.50]	Peripheral-heavy
Net-10	Anna Karenina Net	138	493	2.48	[0.22, 0.10, 0.68]	Peripheral-heavy
Net-11	US Airports ‘97 Net	332	2126	3.22	[0.27, 0.18, 0.55]	Peripheral-heavy
Net-12	EU Air Transport Net	405	1981	3.81	[0.24, 0.21, 0.55]	Peripheral-heavy

scale-free networks with lower λ_{sp} values are core-heavy on their own or in association with an another class; scale-free networks with larger λ_{sp} values are peripheral-heavy. The two airport networks (US and EU) are peripheral-heavy and this is understandable with the presence of several stub airports compared to the hub airports (even though the number of hub airports is not negligible, vindicating the scale-free degree distribution of these networks). We notice the intermediate nodes to be typically nodes that are adjacent to the core nodes and/or the peripheral nodes; they may not be at the center of the network, but may not be at the boundary of the network as well. It is very important to recognize such nodes and categorize them to a separate class rather than strictly following the core-peripheral two-layer model.

Note that the Yifan Hu proportional layout algorithm was run on the real-world networks without any inputs regarding the CIP index values for the nodes. Only the node size and color were set per the CIP index value and the network classification

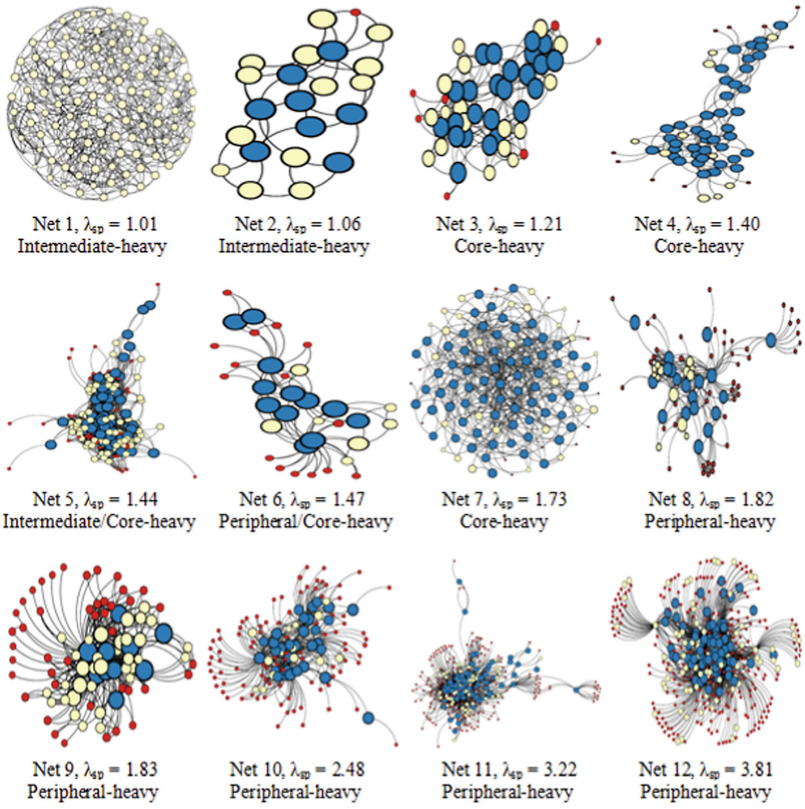


Fig. 5. Yifan Hu Proportional Layout Algorithm-based Visualization of the Real-World Networks [Node Size and Color set per CIP Index Values and Classification]

respectively. The presence of the blue-colored nodes at the core/center of the network and the red-colored nodes at the periphery of the network per the Yifan Hu proportional layout algorithm is an implicit validation of our approach.

4 Related Work

Borgatti and Everett [17] formulated a hub-and-spoke structure for the core-peripheral nodes in a network: the core nodes are mainly adjacent to other core nodes and in some instances adjacent to peripheral nodes; whereas, peripheral nodes do not connect with other peripheral nodes. The above model formulation fits mainly for scale-free networks, but not for random networks. On the other hand, the *k*-core measure [5, 18] divides the network to a layered hierarchy and is by far the most commonly used centrality metric in the literature to capture the coreness of a node. The *k*-core measure is strongly-correlated to the EVC of the nodes; but, it does not take into account the participation of the nodes in the shortest paths between any two nodes as well as the proximity of the nodes to the rest of the nodes in the network. The *k*-core measure is vulnerable (like in the case of the

running example graph of Sect. 2) to find a highly dense sub graph (for a larger k value) that may be farther away from the rest of the nodes in the network. Besides, the k-core value for a node will be just an integer and several nodes could incur the same value of k. Hence, the k-core procedure could lead to ambiguity in the ranking of the nodes (two or more nodes might incur the same k-core value). In 2010, a centrality metric referred to as coreness [1] of a node was formulated as the sum of the k-core measures of the neighbors of the node. Garas et al. [18] define the weighted degree of a node as the sum of its degree and the weights of the edges incident on the node. Proposals (e.g., [19, 20]) that weigh these two terms are also available in the literature.

The contribution of non-core nodes (the intermediate nodes and peripheral nodes in our jargon) has been recognized in [21] for the effective spread of social media protests from the core nodes (forming the epicenter) to the rest of the nodes in the network; though the non-core nodes may not have as many neighbors as the core nodes, the sheer number of such non-core nodes (especially for scale-free networks) was stated to be critical for information diffusion. In [1], the authors state that the most influential spreaders in a social network may not be the most central nodes with high BWC; thus, acknowledging the contribution of the non-core nodes for information diffusion and epidemic spread. In [2, 3], the authors advocated a core-periphery structure for networks wherein there are multiple cores and multiple peripheries, rather than a single hub-and-spoke model. This in turn implies that not all cores and peripheries are of the same size and density, indicating a need for a third class of nodes (the intermediate nodes).

5 Conclusions and Future Work

Our contributions in this paper are the following: (1) We propose a quantitative measure (referred to as Core-Intermediate-Peripheral: CIP Index) to capture the extent with which a node could serve as core node in a network. Unless two nodes incur the same values for all the four centrality metrics (DEG, EVC, BWC, CLC) based on which the CIP index values for the nodes are determined, the CIP index values (a real-valued measure) for any two nodes are expected to be different and could be used to unambiguously rank the nodes in the network with respect to the extent the nodes could serve as core nodes. (2) Rather than two classes (core or peripheral), we propose three classes of nodes (core, intermediate and peripheral) and the CIP index value incurred for a node could be used to assign the class to which a node belongs to. The 3-element CIP bins-fraction tuple determined for a network could be used to decide whether the network is heavy with respect to a particular class or two of the three classes. (3) Upon evaluation of the centrality metrics datasets of a suite of 12 diverse real-world networks, we observe random networks to be intermediate-heavy; scale-free networks with lower variation in node degree to be core-heavy on their own or in association with an another class (intermediate or peripheral) and scale-free networks with relatively larger variation in node degree to be peripheral-heavy. As part of future work, we plan to conduct factor analysis on centrality metrics datasets of real-world networks with the above four centrality metrics as features and determine the number of latent factors behind the centrality metric values incurred by nodes in random networks and scale-free networks with lower and higher variations in node degree.

Acknowledgement. The work leading to this paper was partly funded through the U.S. National Science Foundation (NSF) grant OAC-1835439 and partly supported through a sub contract received from University of Virginia titled: Global Pervasive Computational Epidemiology, with the National Science Foundation as the primary funding agency. The views and conclusions contained in this paper are those of the authors and do not represent the official policies, either expressed or implied, of the funding agency.

References

1. Kitsak, M., et al.: Identification of influential spreaders in complex networks. Nat. Phys. **6**(11), 888–893 (2010)
2. Rombach, M.P., Porter, M.A., Fowler, J.H., Mucha, P.J.: Core-periphery structure in networks. SIAM J. Appl. Math. **74**(1), 167–190 (2014)
3. Kojaku, S., Masuda, N.: Finding multiple core-peripheral pairs in networks. Phys. Rev. E **96**(052313), 1–19 (2017)
4. Cormen, T.H., Leiserson, C.E., Rivest, R.L., Stein, C.: Introduction to Algorithms, 3rd edn. MIT Press, Cambridge (2022)
5. Newman, M.E.J.: Networks: An Introduction, 1st edn. Oxford University Press, Oxford (2010)
6. Bonacich, P.: Power and centrality: a family of measures. Am. J. Sociol. **92**(5), 1170–1182 (1987)
7. Freeman, L.: A set of measures of centrality based on betweenness. Sociometry **40**(1), 35–41 (1977)
8. Freeman, L.: Centrality in social networks: conceptual classification. Soc. Netw. **1**(3), 215–239 (1979)
9. Mulaik, S.A.: Foundations of Factor Analysis, 2nd edn. Chapman & Hall, London (2009)
10. Strang, G.: Linear Algebra and its Applications, 4th edn. Brooks Cole, Pacific Grove (2006)
11. https://pypi.org/project/factor-analyzer/
12. Erdos, P., Renyi, A.: On random graphs I. Publicationes Mathematicae **6**(3–4), 290–297 (1959)
13. Albert, R., Barabasi, A.-L.: Statistical mechanics of complex networks. Rev. Mod. Phys. **74**(47), 47–97 (2002)
14. https://gephi.org/tutorials/gephi-tutorial-layouts.pdf
15. Meghanathan, N.: Spectral radius as a measure of variation in node degree for complex network graphs. In: Proceedings of the 3rd International Conference on Digital Contents and Applications, Hainan, China, 20–23 December 2014, pp. 30–33 (2014)
16. https://gephi.org/
17. Borgatti, S.P., Everett, M.G.: Models of core/periphery structures. Soc. Netw. **21**(4), 375–395 (2000)
18. Garas, A., Schweitzer, F., Havlin, S.: A K-shell decomposition method for weighted networks. New J. Phys. **14**(8), 083030 (2012)
19. Eidsaa, M., Almaas, E.: S-core network decomposition: a generalization of k-core analysis to weighted networks. Phys. Rev. E **88**(6), 062819 (2013)
20. Wei, B., Liu, J., Wei, D., Gao, C., Deng, Y.: Weighted K-shell decomposition for complex networks based on potential edge weights. Physica A **420**, 277–283 (2015)
21. Barbera, P., et al.: The critical periphery in the growth of social protests. PLoS ONE **0143611**, 1–15 (2015)

Multipole Engineering of Optical Forces

Denis Kislov[1] and Vjaceslavs Bobrovs[2(✉)]

[1] Center for Photonics and 2D Materials, Moscow Institute of Physics and Technology, Dolgoprudny 141700, Russia
[2] Institute of Telecommunications, Riga Technical University, Riga 1048, Latvia
vjaceslavs.bobrovs@rtu.lv

Abstract. In most cases, the smallest particles are subjected to forces of attraction and repulsion due to the gradient and radiation pressure. The use of additional degrees of freedom in optomechanical manipulations opens up new possibilities. In this work, we investigate the optical forces acting on a high refractive index silicon sphere in a focused Gaussian beam. Analysis of multipolarity allows us to separate the optical force into trapping and scattering components, which have different physical nature and result in different mechanical effects. These effects are associated with multipolar interference and asymmetric far-field scattering patterns. Such an approach is useful for understanding optomechanical motion and its regularities in the far-field zone. Multipolar engineering of optical forces, not limited to spherical shapes, opens up new possibilities in microfluidic applications, including sorting and assembly of complex volumetric geometries. #COMESYSO1120.

Keywords: optical forces · multipole expansion · gaussian beam · optical trapping · optical bending · optical antitrapping

1 Introduction

The light can exert a mechanical influence on nanoparticles, causing them to move [1]. The high refractive index of the material from which these particles are made allows them to efficiently interact with light and induce mechanical effects [2–5].

One such effect is optical thermal interaction, where the light radiation heats the nanoparticles, resulting in their mechanical movement - a phenomenon known as thermophoretic forces [6–9]. This can be used, for example, in the creation of micro or nanomotors, where light is used to move nanoparticles [10, 11].

Another important effect is the optical forces acting on the nanoparticles from the light. When the nanoparticle absorbs or scatters light, the change in light momentum leads to the emergence of a force that can cause mechanical movement or deformation of the nanoparticle. Optical forces can also arise in the presence of gradients of the electromagnetic field in space - these are called gradient optical forces [12, 13].

Additionally, nanoparticles with a high refractive index can exhibit mechanical resonances, meaning they can vibrate at certain frequencies when exposed to light. This can be utilized in the development of sensitive sensor devices or optical amplifiers.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 373–379, 2024.
https://doi.org/10.1007/978-3-031-53552-9_33

Thus, the mechanical interaction of high refractive index nanoparticles with light [14–16] is an area of research that explores the effects arising from the interaction of light with the movements [17–23], vibrations, or deformations of nanoparticles based on their high refractive index. This can be important for the development of new optical and mechanical devices [24–27] and systems [28].

In this case, it is convenient to conduct an analytical description of optomechanics of high refractive index nanoparticles in terms of the multipole expansion of the electromagnetic field scattered by a nanoantenna [29–31].

2 Methods

In our study, a spherical silicon particle is placed in the focal region of a linearly polarized Gaussian beam. The forces will be analyzed in the focal plane, perpendicular to the propagation direction (Fig. 1). Special attention will be given to the analysis of far-field scattering patterns, which will determine the control conditions for trapping/anti-trapping.

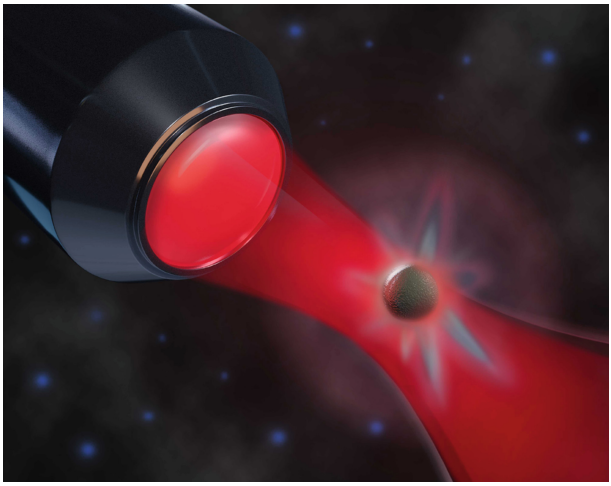


Fig. 1. The geometry of the problem. Silicon particle in a Gaussian beam with w_0 beam waist.

Time-averaged optical force ($\langle \mathbf{F} \rangle$) is:

$$\langle \mathbf{F} \rangle = \frac{1}{2}\mathrm{Re} \oint_S \left[\varepsilon_0 \mathbf{E} \otimes \mathbf{E}^* + \mu_0 \mathbf{H} \otimes \mathbf{H}^* - \frac{1}{2}\left(\varepsilon_0 \mathbf{E} \cdot \mathbf{E}^* + \mu_0 \mathbf{H} \cdot \mathbf{H}^*\right) \overleftrightarrow{\mathbf{I}} \right] \cdot \mathbf{n} dS, \quad (1)$$

where $\mathbf{n}$ is a normal vector.

Where $\mathbf{E} = \mathbf{E}_{inc} + \mathbf{E}_{sca}$, $\mathbf{H} = \mathbf{H}_{inc} + \mathbf{H}_{sca}$ the fields represent fully self-consistent electromagnetic fields, decomposed into the sum of incident and scattered contributions. A integration sphere with a radius that satisfies far-field conditions at its boundary will

be chosen for convenience. In this case, the total force can be decomposed into two contributions:

$$\left\langle {}^{I}\mathbf{F}\right\rangle = \lim_{r\to\infty}\oint_{S}\mathrm{Re}\left[-\frac{1}{2}\left(\varepsilon_0\mathbf{E}_{inc}\cdot\mathbf{E}^{*}_{sca}+\mu_0\mathbf{H}_{inc}\cdot\mathbf{H}^{*}_{sca}\right)\mathbf{n}\right]dS, \tag{2}$$

$$\begin{aligned}\left\langle {}^{R}\mathbf{F}\right\rangle &= -\frac{1}{4}\lim_{r\to\infty}\oint_{S}\mathrm{Re}\left[\left(\varepsilon_0\mathbf{E}_{sca}\cdot\mathbf{E}^{*}_{sca}+\mu_0\mathbf{H}_{sca}\cdot\mathbf{H}^{*}_{sca}\right)\cdot\mathbf{n}\right]dS\\ &= -\frac{1}{2}\lim_{r\to\infty}\oint_{S}\mathrm{Re}\left[\left(\varepsilon_0\mathbf{E}_{sca}\cdot\mathbf{E}^{*}_{sca}\right)\cdot\mathbf{n}\right]dS\end{aligned}$$

where $\left\langle {}^{I}\mathbf{F}\right\rangle$ (interception or extinction force) emerges from the interference between incident and scattered fieds, while $\left\langle {}^{R}\mathbf{F}\right\rangle$ (recoil force) depends only on scattering. The later is governed by multipolar interference, as it was shown in [32–35]. The recoil force can be directly linked to the scattering diagram asymmetry [36]. Multipolar decomposition of the far field is given by [37, 38]:

$$\begin{cases}\mathbf{E}_{sca}(\mathbf{r}) = \frac{k^2}{4\pi\varepsilon_0}\frac{e^{ikr_0}}{r_0}\left[\mathbf{n}\times(\mathbf{p}\times\mathbf{n})+\frac{1}{c}(\mathbf{m}\times\mathbf{n})+\frac{ik}{2}\mathbf{n}\times\left[\mathbf{n}\times(\mathbf{Q}^{e}\cdot\mathbf{n})\right]+\frac{ik}{2c}\mathbf{n}\times(\mathbf{Q}^{m}\cdot\mathbf{n})\right],\\ \mathbf{H}_{sca}(\mathbf{r}) = \frac{1}{\mu_0}\mathbf{B}_{sca}(\mathbf{r}) = \frac{1}{Z}\mathbf{n}\times\mathbf{E}_{sca}(\mathbf{r})\end{cases} \tag{3}$$

where Z is the wave impedance. Multipolar moments are given by:

$$\mathbf{p} = \varepsilon_0\alpha_e\mathbf{E}_{inc},\quad \mathbf{m} = \frac{\alpha_m}{\mu_0}\mathbf{B}_{inc},\quad \mathbf{Q}^{e} = \varepsilon_0\alpha_{Q^e}\frac{\nabla\mathbf{E}_{inc}+\mathbf{E}_{inc}\nabla}{2},\quad \mathbf{Q}^{m} = \frac{\alpha_{Q^m}}{\mu_0}\frac{\nabla\mathbf{B}_{inc}+\mathbf{B}_{inc}\nabla}{2} \tag{4}$$

where the relation $(\nabla\mathbf{A}+\mathbf{A}\nabla)_{ij} = \partial_i A_j + \partial_j A_i$ is used. Polarizabilities are linked to Mie coefficients (a_1, a_2, b_1, and b_2) as follows:

$$\alpha_e = i\frac{6\pi}{k^3}a_1,\quad \alpha_m = i\frac{6\pi}{k^3}b_1,\quad \alpha_{Q^e} = i\frac{40\pi}{k^5}a_2,\quad \alpha_{Q^m} = i\frac{40\pi}{k^5}b_2 \tag{5}$$

3 Results

By substituting Eq. 3 into Eq. 2 the following force decomposition is obtained:

$$\begin{cases}\langle F_i\rangle = \left\langle {}^{I}F_i\right\rangle + \left\langle {}^{R}F_i\right\rangle\\ \left\langle {}^{I}F_i\right\rangle = \frac{1}{2}\mathrm{Re}\left[p_j\nabla_i E^{*}_{inc,j}\right]+\frac{1}{2}\mathrm{Re}\left[m_j\nabla_i B^{*}_{inc,j}\right]+\frac{1}{4}\mathrm{Re}\left[(Q^{e})_{jk}\nabla_i\nabla_k E^{*}_{inc,j}\right]\\ +\frac{1}{4}\mathrm{Re}\left[(Q^{m})_{jk}\nabla_i\nabla_k B^{*}_{inc,j}\right]\\ \left\langle {}^{R}F_i\right\rangle = -\frac{k^4}{12\pi\varepsilon_0 c}\mathrm{Re}\left[\varepsilon_{ijk}p_j m^{*}_k\right]-\frac{k^5}{40\pi\varepsilon_0}\mathrm{Im}\left[(Q^{e})_{ij}p^{*}_j\right]-\frac{k^5}{40\pi\varepsilon_0 c^2}\mathrm{Im}\left[(Q^{m})_{ij}m^{*}_j\right]\\ -\frac{k^6}{240\pi\varepsilon_0 c}\mathrm{Re}\left[\varepsilon_{ijk}(Q^{e})_{lj}(Q^{m})^{*}_{lk}\right],\quad i,j,k \text{ can be x,y or z}\end{cases} \tag{6}$$

Equation 6 incorporates self-consistent electromagnetic fields and their spatial derivatives. It is worth noting that these expressions include cross-terms that arise from the interference of multipole contributions in the far-field region.

It is worth noting that in this study, we utilize spherical multipole moments as the foundation. In this approach, the toroidal moments, which typically require explicit inclusion in a Cartesian expansion, are inherently incorporated.

Although Eq. 6 provides a formalism for calculating optical forces, it can be further adjusted and transformed into a form that explicitly demonstrates the connection between the asymmetry of the scattering pattern and the recoil forces. Through certain mathematical manipulations, the component of the recoil force can be expressed as:

$$\begin{aligned}\left\langle {}^{R}F\right\rangle = &-\frac{1}{c}\oint\limits_{S}\left(I_p + I_m + I_{Q^e} + I_{Q^m} + c\varepsilon_0 \mathrm{Re}\left[E_{Q^e}\cdot E_m^* + E_{Q^m}\cdot E_p^*\right]\right)\cdot ndS - \\ &-\oint\limits_{S}\left(\varepsilon_0 \mathrm{Re}\left[E_p\cdot E_m^* + E_{Q^e}\cdot E_{Q^m}^* + E_{Q^e}\cdot E_p^* + E_{Q^m}\cdot E_m^*\right]\right)\cdot ndS,\end{aligned} \quad (7)$$

where far-field multipoles intensities $I_p,\ I_m,\ I_{Q^e},\ I_{Q^m}$. The electric fields generated by multipole moments can exhibit either symmetric or antisymmetric characteristics. As a result, when multiplying fields of different symmetries, angular asymmetry emerges in the cross-terms of Eq. 7. Taking this into account, combined with the fact that individual multipoles have a symmetric intensity distribution, the first integral in Eq. 7 becomes zero, resulting in a simplified expression for the recoil force.

$$\left\langle {}^{R}\mathbf{F}\right\rangle = -\oint\limits_{S}\left(\varepsilon_0 \mathrm{Re}\left[\mathbf{E}_p\cdot\mathbf{E}_m^* + \mathbf{E}_{Q^e}\cdot\mathbf{E}_{Q^m}^* + \mathbf{E}_{Q^e}\cdot\mathbf{E}_p^* + \mathbf{E}_{Q^m}\cdot\mathbf{E}_m^*\right]\right)\cdot\mathbf{n}dS \quad (8)$$

To determine the transverse components of the recoil optical force, the vectorial force can be projected onto the Cartesian coordinate system using the following method:

$$\begin{aligned}
&{}^{R}F_x^{pm} = -\tfrac{k^4}{12\pi\varepsilon_0 c}\mathrm{Re}\left[p_y m_z^* - p_z m_y^*\right];\quad {}^{R}F_y^{pm} = -\tfrac{k^4}{12\pi\varepsilon_0 c}\mathrm{Re}\left[p_z m_x^* - p_x m_z^*\right]\\
&{}^{R}F_x^{Q^e Q^m} = -\tfrac{k^6}{240\pi\varepsilon_0 c}\mathrm{Re}\left[\begin{array}{l}(Q^e)_{xy}(Q^m)_{xz}^* - (Q^e)_{xz}(Q^m)_{xy}^* + (Q^e)_{yy}(Q^m)_{yz}^* - (Q^e)_{yz}(Q^m)_{yy}^* + \\ +(Q^e)_{zy}(Q^m)_{zz}^* - (Q^e)_{zz}(Q^m)_{zy}^*\end{array}\right]\\
&{}^{R}F_y^{Q^e Q^m} = -\tfrac{k^6}{240\pi\varepsilon_0 c}\mathrm{Re}\left[\begin{array}{l}-(Q^e)_{xx}(Q^m)_{xz}^* + (Q^e)_{xz}(Q^m)_{xx}^* - (Q^e)_{yx}(Q^m)_{yz}^* + (Q^e)_{yz}(Q^m)_{yx}^* - \\ -(Q^e)_{zx}(Q^m)_{zz}^* + (Q^e)_{zz}(Q^m)_{zx}^*\end{array}\right]\\
&{}^{R}F_x^{Q^e p} = -\tfrac{k^5}{40\pi\varepsilon_0}\mathrm{Im}\left[(Q^e)_{xx}p_x^* + (Q^e)_{xy}p_y^* + (Q^e)_{xz}p_z^*\right]\\
&{}^{R}F_y^{Q^e p} = -\tfrac{k^5}{40\pi\varepsilon_0}\mathrm{Im}\left[(Q^e)_{yx}p_x^* + (Q^e)_{yy}p_y^* + (Q^e)_{yz}p_z^*\right]\\
&{}^{R}F_x^{Q^m m} = -\tfrac{k^5}{40\pi\varepsilon_0}\mathrm{Im}\left[(Q^m)_{xx}m_x^* + (Q^m)_{xy}m_y^* + (Q^m)_{xz}m_z^*\right]\\
&{}^{R}F_y^{Q^m m} = -\tfrac{k^5}{40\pi\varepsilon_0}\mathrm{Im}\left[(Q^m)_{yx}m_x^* + (Q^m)_{yy}m_y^* + (Q^m)_{yz}m_z^*\right]
\end{aligned} \quad (9)$$

Equation 9 provides explicit expressions underlining the link between the scattered pattern asymmetry and attraction/repulsion recoil forces.

4 Discussions

It is important to note the following: while the constructive interference between the terms in Eq. 6 has a minimal impact on the behavior of the force, destructive interference does occur and results in a complete cancellation of the overall force. This cancellation occurs even in the presence of a moderately high gradient of field intensity. The total

optical force equals zero only for certain wavelengths at $\theta = 0$ and $\theta = \pi / 2$ (as shown in Fig. 2). Additionally, there are regions where the net force approaches zero for all angles, although not exactly equal to zero. These regions correspond to wavelengths of 600, 625, and 750 nm (as indicated by the spectral-angular dependencies of the contours where net, interception, and recoil forces are zero in Fig. 2).

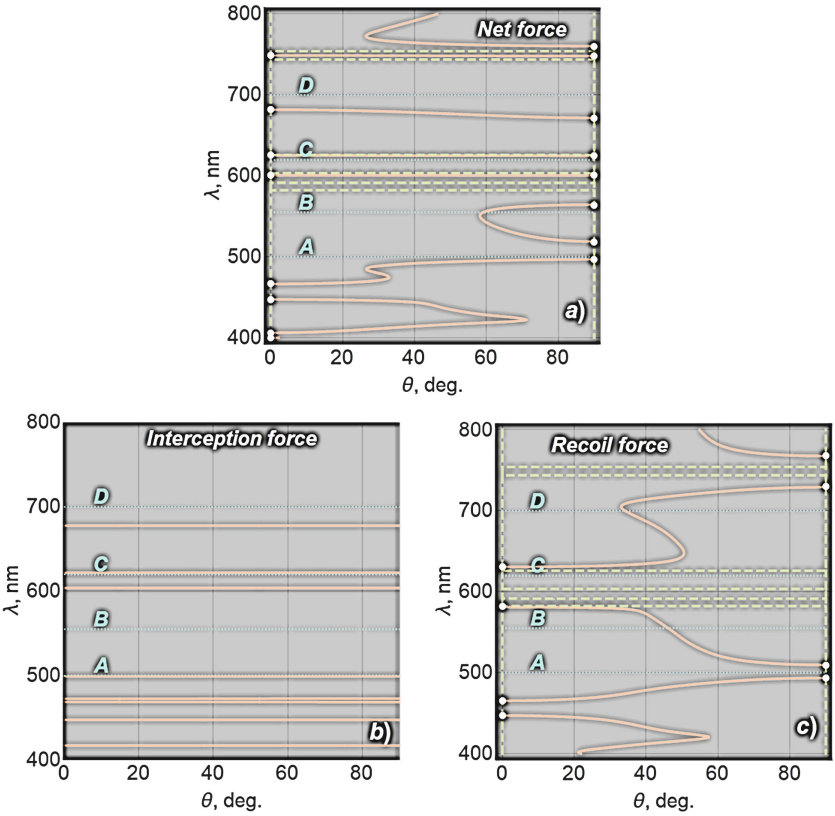


Fig. 2. The contours depicted in the figure illustrate the regions where the radial (orange curves) and azimuthal (green curves) components of the net optical force have values of zero. The white circles indicate points where both components simultaneously equal zero.

5 Conclusion

The study focused on investigating the influence of higher-order multipoles on the interaction between focused laser beams and high refractive index particles. It was demonstrated that by carefully managing the contributions of different multipolar components, the direction of optical forces can be controlled.

In addition, it was observed that the interception force acting on a high refractive index particle can change its sign depending on the wavelength of the incident

light, resulting in an anti-trapping regime. This behavior deviates from the conventional trapping regime typically observed in classical optical tweezers.

The incorporation of multipolar degrees of freedom into optomechanical interactions expands the potential for precise motion control at the nanoscale, offering new possibilities for applications in optics, biology, medicine, and lab-on-a-chip platforms.

Acknowledgements. The authors gratefully acknowledge the financial support from the Ministry of Science and Higher Education of the Russian Federation (Agreement No. № 075-15-2022-1150). V.B. acknowledges the support of the Latvian Council of Science, project: PHOTON, No. lzp-2022/1-0579.

References

1. Ashkin, A., Dziedzic, J.M., Bjorkholm, J.E., Chu, S.: Observation of a single-beam gradient force optical trap for dielectric particles. Opt. Lett. **11**, 288 (1986)
2. Terekhov, P.D., Evlyukhin, A.B., Redka, D., Volkov, V.S., Shalin, A.S., Karabchevsky, A.: Magnetic octupole response of dielectric quadrumers. Laser Photon. Rev. **14**(4), 1900331 (2020)
3. Barhom, H., et al.: Biological kerker effect boosts light collection efficiency in plants. Nano Lett. **19**(10), 7062 (2019)
4. Canós Valero, A., Gurvitz, E.A., Benimetskiy, F.A., et al.: Theory, observation, and ultrafast response of the hybrid anapole regime in light scattering. Laser Photon. Rev. **15**(10), 2100114 (2021)
5. Vestler, D., et al.: Circular dichroism enhancement in plasmonic nanorod metamaterials. Opt. Exp. **26**(14), 17841–17848 (2018)
6. Würger, A.: Thermal non-equilibrium transport in colloids. Rep. Prog. Phys. **73**(12), 126601 (2010)
7. Piazza, R., Parola, A.: Thermophoresis in colloidal suspensions. J. Phys. Condens. Matter **20**(15), 153102 (2008)
8. Lin, L., et al.: Opto-thermoelectric pulling of light-absorbing particles. Light Sci. Appl. **9**(1), 34 (2020)
9. Zhang, Y., et al.: Optical attraction of strongly absorbing particles in liquids. Opt. Exp. **27**(9), 12414–12423 (2019)
10. Gargiulo, J., et al.: Understanding and reducing photother-mal forces for the fabrication of au nanoparticle dimers by optical printing. Nano Lett. **17**(9), 5747–5755 (2017)
11. Joby, J.P., et al. Optically-assisted thermophoretic reversible assembly of colloidal particles and E. Coli using graphene oxide microstructures. Sci. Rep. **12**(1), 3657 (2022)
12. Albaladejo, S., Marqués, M.I., Laroche, M., Sáenz, J.J.: Scattering forces from the curl of the spin angular momentum of a light field. Phys. Rev. Lett. **102**, 113602 (2009)
13. Kostina, N.A., Kislov, D.A., Ivinskaya, A.N., et al.: Nanoscale tunable optical binding mediated by hyperbolic metamaterials. ACS Photonics **7**(2), 425–433 (2020)
14. Shalin, A.S., Sukhov, S.V.: Plasmonic nanostructures as accelerators for nanoparticles: optical nanocannon. Plasmonics **8**(2), 625–629 (2013)
15. Simovski, C.R., Shalin, A.S., Voroshilov, P.M., Belov, P.A.: Photovoltaic absorption enhancement in thin-film solar cells by non-resonant beam collimation by submicron dielectric particles. J. Appl. Phys. **114**, 103104 (2013)
16. Kucherik, A., et al.: Nano-antennas based on silicon-gold nanostructures. Sci. Rep. **9**, 338 (2019)

17. Izmodenova, S.V., Kislov, D.A., Kucherenko, M.G.: Accelerated nonradiative electron-excitation energy transfer between molecules in aqueous pools of reverse micelles containing encapsulated silver nanoparticles. Colloid J. **76**, 683 (2014)
18. Kislov, D.: Effect of plasmonic silver nanoparticles on the photovoltaic properties of graetzel solar cells. Phys. Procedia **73**, 114 (2015)
19. Kucherenko, M.G., Kislov, D.A.: Plasmon-activated intermolecular nonradiative energy transfer in spherical nanoreactors. J. Photochem. Photobiol. A Chem. **354**, 25 (2018)
20. Kislov, D., Novitsky, D., Kadochkin, A., Redka, D., Shalin, A.S., Ginzburg, P.: Diffusion-inspired time-varying phosphorescent decay in a nanostructured environment. Phys. Rev. B **101**, 035420 (2020)
21. Novitsky, D.V., Valero, A.C., Krotov, A., Salgals, T., Shalin, A.S., Novitsky, A.V.: CPA-lasing associated with the quasibound states in the continuum in asymmetric non-Hermitian structures. ACS Photonics (2022)
22. Shalin, A.S.: Broadband blooming of a medium modified by an incorporated layer of nanocavities. JETP Lett. **91**, 636–642 (2010)
23. Shalin, A.S., Moiseev, S.G.: Optical properties of nanostructured layers on the surface of an underlying medium. Opt. Spectrosc. **106**, 916–925 (2009)
24. Kuznetsov, A.V., Canós Valero, A., Tarkhov, M., Bobrovs, V., Redka, D., Shalin, A.S.: Transparent hybrid anapole metasurfaces with negligible electromagnetic coupling for phase engineering. Nanophotonics **10**(17), 4385–4398 (2021)
25. Novitsky, D., Karabchevsky, A., Lavrinenko, A., Shalin, A., Novitsky, A.: PT symmetry breaking in multilayers with resonant loss and gain locks light propagation direction. Phys. Rev. B **98**, 125102 (2018)
26. Novitsky, D.V., Shalin, A.S., Redka, D., Bobrovs, V., Novitsky, A.V.: Quasibound states in the continuum induced by PT symmetry breaking. Phys. Rev. B **104**(8), 85126 (2021)
27. Shamkhi, H.K., Sayanskiy, A., Valero, A.C., et al.: Transparency and perfect absorption of all-dielectric resonant metasurfaces governed by the transverse Kerker effect. Phys Rev Mater **3**(8), 85201 (2019)
28. Canós Valero, A., Kislov, D., Gurvitz, E.A., et al.: Nanovortex-driven all-dielectric optical diffusion boosting and sorting concept for lab-on-a-chip platforms. Adv. Sci. **7**(11), 1903049 (2020)
29. Kislov, D., et al.: Multipole engineering of attractive-repulsive and bending optical forces. Adv. Photonics Res. **2**, 2100082 (2021)
30. Kuznetsov, A.V., et al.: Special scattering regimes for conical all-dielectric nanoparticles. Sci. Rep. **12**, 21904 (2022)
31. Terekhov, P.D., Evlyukhin, A.B., Shalin, A.S., Karabchevsky, A.: Polarization-dependent asymmetric light scattering by silicon nanopyramids and their multipoles resonances. J. Appl. Phys. **125**, 173108 (2019)
32. Salandrino, A., Fardad, S., Christodoulides, D.N.: Generalized Mie theory of optical forces. J. Opt. Soc. Am. B **29**, 855 (2012)
33. Duan, X.Y., Wang, Z.G.: Fano resonances in the optical scattering force upon a high-index dielectric nanoparticle Phys. Rev. A **96**, 053811 (2017)
34. Chen, J., Ng, J., Lin, Z., Chan, C.T.: Optical pulling force. Nat. Photon **5**, 531–534 (2011)
35. Nieto-Vesperinas, M., Sáenz, J.J., Gómez-Medina, R., Chantada, L.: Optical forces on small magnetodielectric particles. Opt. Exp. **18**, 11428–11443 (2010)
36. Kiselev, A., Achouri, K., Martin, O.J.F.: Multipole interplay controls optical forces and ultra-directional scattering. Opt. Exp. **28**, 27547 (2020)
37. Raab, R.E.: Multipole Theory in Electromagnetism (2005)
38. Bohren, C.F., Huffman, D.R.: Absorption and Scattering of Light by Small Particles. Wiley (1998)

Data Mart in Business Intelligence with Ralph Kimball for Commercial Sales

Alessandro Chanco Torres(✉), Angel Quiñonez Gastelu, Juan J. Soria, Mercedes Vega Manrique, and Lidia Segura Peña

Universidad Tecnológica del Perú, Lima, Peru
u18200318@utp.edu.pe

Abstract. Business intelligence in organizations, not exploiting the volume of their data is a problem, if they do not have technological tools that allow real-time reporting. This article presents a well-executed study on the implementation of a Data Mart in an organization, which has resulted in a remarkable increase in sales. It effectively leverages Ralph Kimball's methodology and the star model, providing rich insight to decision makers. The research is comprehensive, valuable and could greatly benefit other organizations wishing to improve their strategic planning. Extract, Transform, Load (ETL) processes were applied for information gathering and approaches provided by the application of the Ralph Kimball methodology. The Tableau platform was used for sales reports in the areas of warehouse, management and accounting, obtaining dynamic dashboards, showing a 30% increase in sales of the categories, being the most benefited the category of small bread, which went from 42,390 units sold to 59,760 units, representing an increase of 40% sold, there was an increase of 45% in the product pan croissant, as well as an increase of more than 30% in sales in general; all these results are the comparison made from April to May 2023. It is concluded that the Data Mart is efficient in reporting with an optimum of 20% in the organization's products.

Keywords: Business Intelligence · Dashboard · Data mart · ETL Processes · Ralph Kimball

1 Introduction

Globally, the sales process with large volumes of data and information is a big problem because data analysis has been evolving rapidly and is constantly growing, suggesting companies to learn the fundamental skills to use Big Data. In 2000, about 800,000 petabytes were housed in the world, It is expected that by 2020 they will reach 35 zettabytes [1]. In Latin America, most SMEs have great difficulties in the implementation of ICT, in Colombia, SMEs present data in both structured and unstructured form, which leads to the fact that they cannot acquire systems that allow them to extract and use the real value of their data and information, showing problems such as the inaccuracy of not having data from the customers who most need to be informed which products are most purchased, which products are most distributed, and which services are offered [2]. In Peru, one of the main challenges faced by the retail sector is the need to obtain information

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 380–396, 2024.
https://doi.org/10.1007/978-3-031-53552-9_34

for decision making, since the volume of data generated is greater compared to other sectors. In addition, they do not have solutions that implement Business Intelligence and Analytics, which does not allow them to have a vision of the company or to have timely information to support their decision making [3]. In the company Wilton's S.A.C. dedicated to commercial sales, the main reason that its sales do not increase is the disorganization of its information, the lack of use of technologies that help analyze data based on indicators, number of sales, number of products sold, number of customers and products produced, and that is reflected in the incorrect decision making in its sales process.

Companies grow due to the demand for products by customers, as a result, the information of services and / or products provided increases, giving way to a total clutter of information. Currently there are several ways to store information, either manually or digitally, through database managers, simple as Microsoft Access or Micro-soft Excel, and complex así, SQL Server Oracle, SQL Server. Despite the existence of these software are not used, generating loss of time and raising costs at the time of requiring them so that the information is not used in its maximum splendor, which is why companies must implement Data Mart, so that their information is extracted by the same organization quickly, efficiently and accurately with a minimum percentage of error and so that the different areas of the company can take advantage of the required information and assist in the decisions taken.

The Business Intelligence tools that companies usually use are Microsoft Excel, BI Power, SQL Server, SAP, Oracle. Through these tools, solutions were delivered that are the deliverables handled by business analysts, presenting dashboards, gathering reports [4].

The business intelligence approach designs an inoperability framework using technologies to integrate data and analytically develop queries of large volumes of data, digitize and innovate technologically in relation to data to continue providing services in a varied way in conjunction with information security and data management [5]. Boukobza hypothesizes that a structured model using clustering methods helps demonstrate quality and continuity of care, supported by expert opinion to present an operative report [6]. T. Thivakaran carried out An Analysis, Design, and Implementation of a Data Mart were conducted to ensure effective decision-making in the sales department of PROMED E.I.R.L. As part of this process, data was collected from the company, and the information from the ETL process and the Data Mart were carefully analyzed. This initiative successfully resolved issues related to reporting delays and challenges. Additionally, it ensured proper decision-making concerning report quality, information security, and data integrity within the sales area. As a result, 32% of employees expressed satisfaction with the annually updated information [7]. In 2018, Medina created a Data Mart to capture academic productivity indicators in Chile. The Universidad Arturo Prat de Chile (UNAP) has an Institutional Analysis Unit (UAI) responsible for managing the Data Mart. The purpose of the Data Mart is to provide qualitative and quantitative reports on the university's management across various areas. Due to the substantial amount of information, they proposed developing a Data Warehousing Process (PDW) specifically focused on obtaining indicators related to academic productivity [8]. In 2019, Dianderas

conducted a comprehensive development of a Sales Data Mart designed for medium-sized companies located in Lima, Peru. This involved creating a Data Mart, performing the ETL (Extract, Transform, Load) process, integrating data, and generating reports using various tools, such as SQL Server Integration Services, SQL Server Analysis Services, Visual Studio 2017, and Excel 2010/Power BI. The Kimball methodology was applied due to the specific nature of the project. The outcomes of this initiative were remarkable, with notable benefits including reduced average time for report generation, enhanced report quality and delivery, and the inclusion of detailed and dynamic graphics to facilitate data interpretation [9]. A. Antunes, E. Cardoso y J. Barateiro, support that the data warehousing system and business intelligence are considered in decision making. These systems use ETL and OLAP for the development of the multidimensional database, facilitating the compression of data obtained in an initial [10]. J. Hoelscher, A. Mortimer, describe that using Tableau in a Datamart provides powerful and flexible data visualization capabilities to effectively analyze and present information, which contributes to improved decision making in organizations [11]. S. El Sappagh, A. Hendawi, A.El Bastawissy, define the ETL process as improving the efficiency and effectiveness of the data extraction, transformation and loading stages. Focusing on optimizing processes and ensuring the quality and integrity of data, taken as a tool by companies to improve their decision making and analysis [12]. H. Zhou, K. Chen. S. Wang, in their developed model establish a dynamic strategy for inventory management and establish the optimal price to avoid loss of sales. Obteniendo como resultado que la cantidad de ventas y la calidad de las unidades tienen impacto en la perdida de ventas en distintas temporadas del negocio [13]. G. Peñafiel, V. Yáñez, K. Guaman et al. analyzed the methodology to develop a data warehouse, demonstrating that the methodology shows a high degree of fulfillment of objectives and satisfaction of managers, thanks to its agile approach and worker participation in determining the specific data for the company's decision making [14]. F. Sanchez Galvez, J. Mendoza Moheno, described that through the star model identified areas for improvement of structure and processes to continue to effectively manage the optimization of organizational design with the implementation of the established changes contributed to the growth and development of the organization [15]. O. Sheta, A. Nour Eldeen, mention how data warehouse technology and OLAP processing in the medical field has had an effective solution when evidenced by dashboards, the benefits of a quick and multidimensional view for fast and optimal decision making [16]. M. Zekri, S. Zahaf, F. Gargouri, describes the improvement of BPIS has a positive impact on performance and flexibility between devices. Taking an approach that focuses an approach that focuses on the decision-making dimension, proposing the use of an ontology to characterize the data warehouse schema and capture the multidimensional information of the offers [17]. L. Ilha, F. Junior, describe how the use of data warehouse technology and business intelligence tools evaluated the management proposal seeking to improve decision making and educational quality. Showing results resulting from the conversion of data into knowledge and the contribution to the development of school strategies [18].

The objective of this research is to improve decision making in the sales process with large volumes of data and information presented by the company, applying ETL processes for the collection of information and the approaches provided by the Ralph

Kimball methodology, in order to compete with other companies in the retail sector, grow economically and a future expansion at the regional level.

2 Materials and Methods

2.1 Project Planning

At this stage, all tasks were identified, such as: scheduling the development of tasks, and planning the use of resources. And firstly, information was collected from the company Wilton's SAC and the workflow of the sales area was organized, secondly, reliable reports were generated, in an effective and concise manner, directed to the different areas of the company for decision making and finally there was a dashboard that allowed to elaborate and optimize the required information and present it in a consistent and dynamic way.

2.2 Ralph Kimball Methodology

The research used the approach of the Ralph Kimball methodology according to F. Castañeda, facilitates the creation of a Data Mart, because its design and construction allows to elaborate in a specific way the different areas of an organization, giving a sketch of a future general model of the system [19].

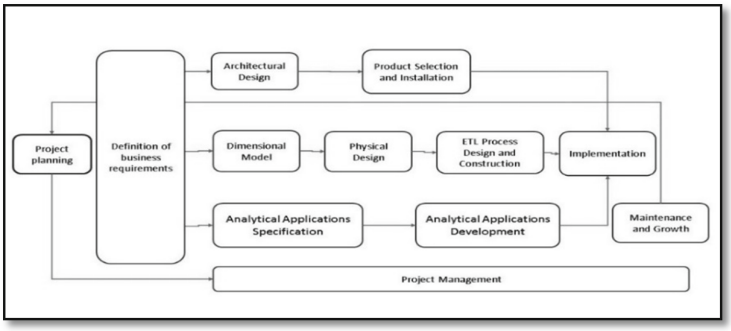


Fig. 1. Ralph Kimball Methodology for Research

The Ralph Kimball methodology, as shown in Fig. 1, was classified into 3 paths, which contain phases that together make up its life cycle. The paths were the technology, which is associated with the design of the architecture of the computer tools used, the data, which allowed the implementation of dimensional models and data transformation, and the application of business intelligence, which allowed the design of logical applications that will show the results of the proposed methodology [19]. In business intelligence there are several recognized methodologies, each one has a different approach; however, we focused on Ralph Kimball, due to its scalability and flexibility, allowing the creation of a Data Mart for the commercial sales area [20].

2.3 Business Model

According to F. Castañeda, the approach of the Ralph Kimball methodology facilitates the creation of a Data Mart, since its design and construction allows it to be developed specifically for the different areas of an organization, obtaining an outline of a future general model of the system [19].

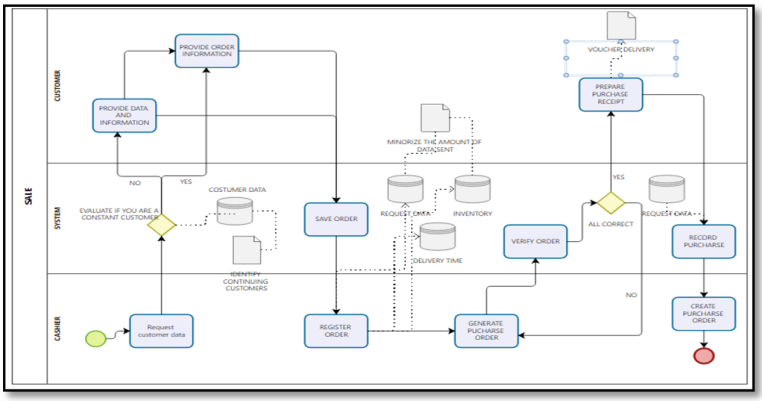


Fig. 2. Business Model of Wilton's SAC developed in Bizagi.

En Fig. 2 shows the flow of the sales process of Wilton's SAC, which shows a data repository called: Customer Data, Order Data, inventory and sales delivery time (Fig. 3).

2.4 Identification of Measurements and the Facts Table

The content of the Facts table will be used by the employees assigned to these areas in different tasks, with the purpose of analyzing and answering the questions presented by the business. These are quantitative data that will go through three processes, filtering, grouping and exploration, having as conditions defined in the Dimensions tables. The Facts tables contain the facts that will be used by DM users to analyze and answer business questions. Through conditions defined in the dimension tables shown in Fig. 4.

2.5 Milestones

In data analysis is a strategic business objective, therefore it is required to plan milestones for reporting in the Tableau platform, which is to determine the stock of the best-selling product in a month, it is also required to determine the best-selling product in a month, as well as to determine o summarize and paraphrase, the purpose is to track both the quantity and value of products sold each month., as well as to determine the total value purchased by a customer in a month, the total value sold by category in a month, the amount of sales of a single product in a month, the amount of sales of a single product in a month and the amount of the least sold product in a month: the total value purchased by a customer in a month, the total value sold by category in a month, the amount of sales of a single product in a month and the stock of the least sold product in a month.

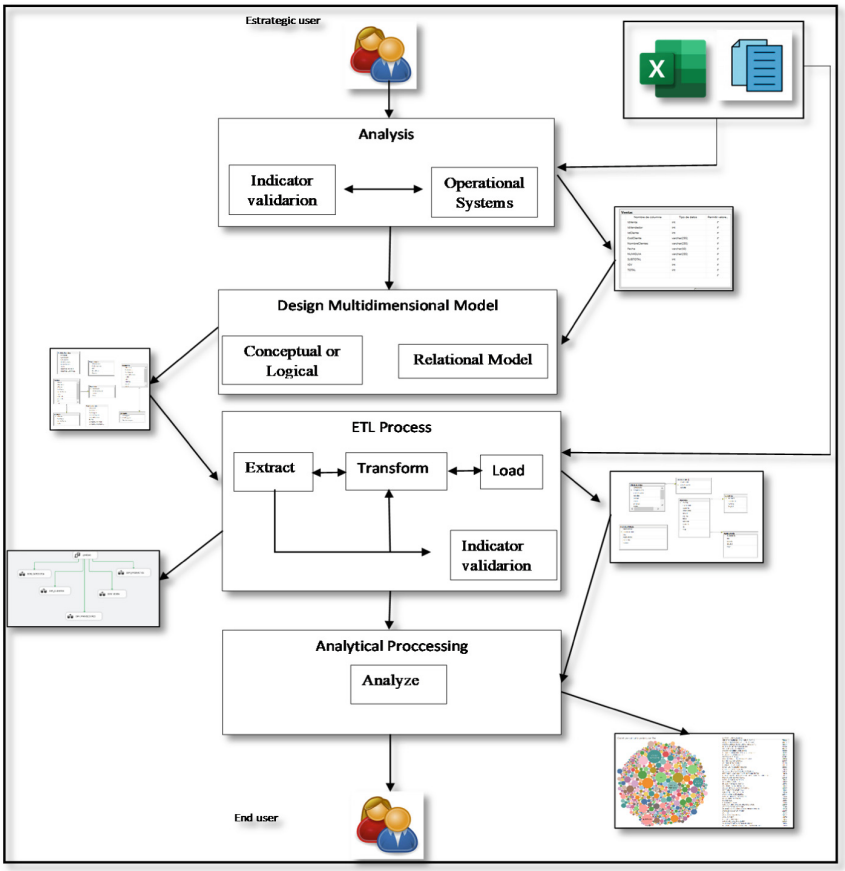


Fig. 3. Research PDM Development Methodology

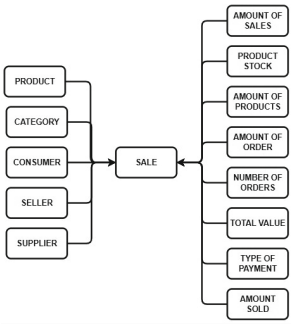


Fig. 4. Dimensions and Milestones Diagram of the research database

2.6 Dimensional Model

In the dimensional modeling, the Perspective Product, Perspective Customer, Perspective Suppliers, Perspective Seller and Perspective Category were designed (Tables 1, 2, 3, 4 and 5).

Table 1. Perspective Product

Perspective Product	
COLUMN NAME	Column Meaning
IDPRODUCT	He is the boss of the client
CODPRODUCT	It is the primary key of the table products
CODCATEGORY	The highlighted category of the product
NAME PRODUCT	It is the name of the product for sale
UNIT	The unit of the product
START	It is the income of product
ENTRANCE	It the amount of product that enters
EXIT	It is how much product was sold
STOCK	It is the amount of product that remains in stock
PLACE	It is the location of the product within the store

Table 2. Perspective Customer

Perspective Customer	
COLUMN NAME	Column Meaning
IDCUSTOMER	It is the primary key of the customers table and is unique for each customer
CODCUSTOMER	It is the customer code
NAMECUSTOMERS	Is the name of the customer
DNI_RUC	It is the client's ID or ruc

2.7 Data Mart Design

The company provides data in.xlsx files that belong to the EXCEL program. By using formulas and tools of the program we filtered and collected the data that helped to exploit the cube with the Tableau platform.

Table 3. Perspective Suppliers

Perspective Suppliers	
COLUMN NAME	Column Meaning
IDSUPPLIER	It is the primary key of the vendors table
CODSUPPLIER	It's the provider's code
RUC	It is the ruc of the company
BUSSINESSNAME	It is the company name
ADDRESS	Is the provider's address
PHONE	It's the provider's phone number

Table 4. Perspective Seller

Perspective Seller	
COLUMN NAME	Column Meaning
IDSELLER	It is the primary key of the Seller table
DNI	It is the ID of the seller
NAMESELLER	It is the name of the seller
USER	It is the user of the seller
PASSWORD	It's the seller's password

Table 5. Perspective Category

Perspective Category	
COLUMN NAME	Column Meaning
IDCATEGORY	It is the primary key of the table and represents a particular category
CODCATEGORY	It is the name of the product category
NAMECATEGORY	The category assigned to the product

2.7.1 Database

Figure 5 shows the database of the company Wilton's SAC without any modification in which there are 6 tables belonging to the organization, the table sales, customers, vendors, suppliers, products and categories, plus 2 additional tables for the results, the table after sales and table after sales success. As can be observed, they do not present an order or a correct connection between them, which shows that there is no technology that supports this database, therefore the database was modified for its correct analysis.

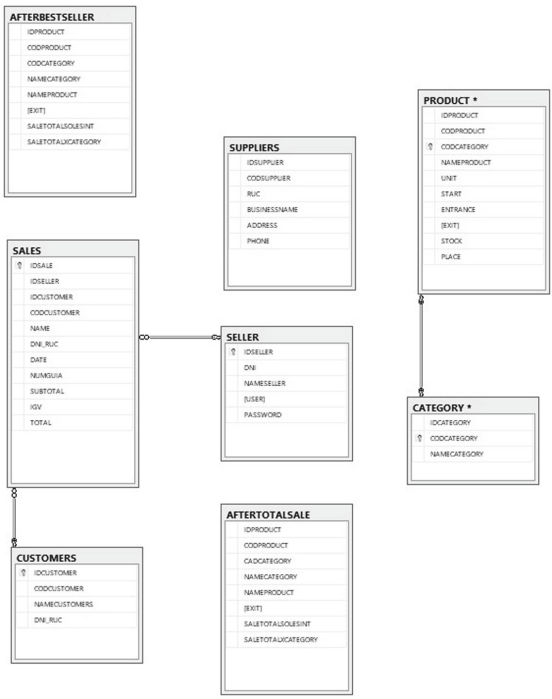


Fig. 5. Wiltons SAC Bakery Database Diagram

2.7.2 Analytical ETL Processing

The ETL process was applied according to A.R. Zaldívar, which consists of the extraction with which all sources of heterogeneous origin were connected, firstly, divided into structured data obtained from relational databases such as SQL Server, My SQL, Oracle, secondly, in semi-structured data are data in files, pdf, XML, Json and finally in unstructured data that has no structure, obtained from reports, audios, videos, web pages. In the transformation, the information obtained was validated and converted into homogeneous data, which were then cleaned, filtered and sorted in a series of necessary operations.

In the data loading, a data model was developed to store all the data used for the analysis, where these data were stored in multidimensional databases or DataWarehouses, from which it was possible to segment the information by areas in the Data Mart. The information was also verified in the target repository. Finally, a Dashboard was created to follow up on the results obtained and to prepare reports that were useful to the organization's managers to make good and better decisions [21].

2.7.3 SQL Server Datamart Design and Development

Figure 6 shows the star model that houses qualities within the organizations, which are based on capabilities, structure, processes, incentives and people [15].

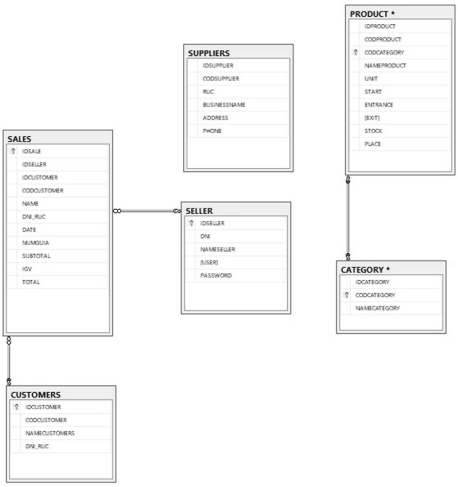


Fig. 6. Design of the star model of the research

2.8 Resolution of Milestones

You want to determine the stock of the best-selling product in a month.

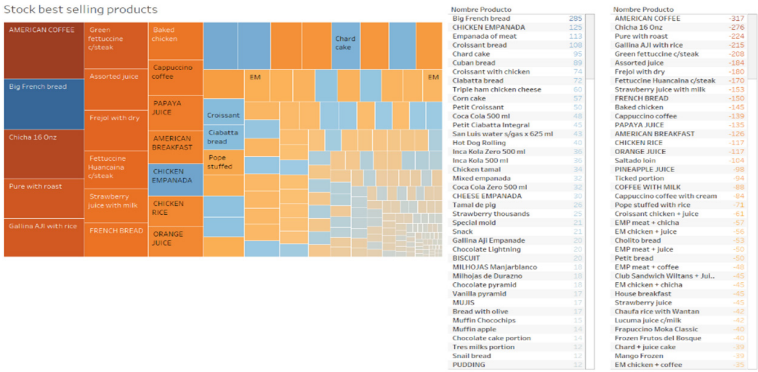


Fig. 7. Stock chart of best-selling products in the month of April.

Figure 7 shows the stock of the most sold products in the month of April, where it can be observed that the main product to run out of stock was american coffee, only 317 units were left in stock, which means that, being the most sold product, it was expected to have enough product for the sales of the month, likewise, the product large French bread, only 285 units were left in stock.

We want to determine the total value sold per category in a month.

La Fig. 8 shows the total value sold by each category in April, with the category of small breads being the most requested and sold, in first place with 42,390 units,

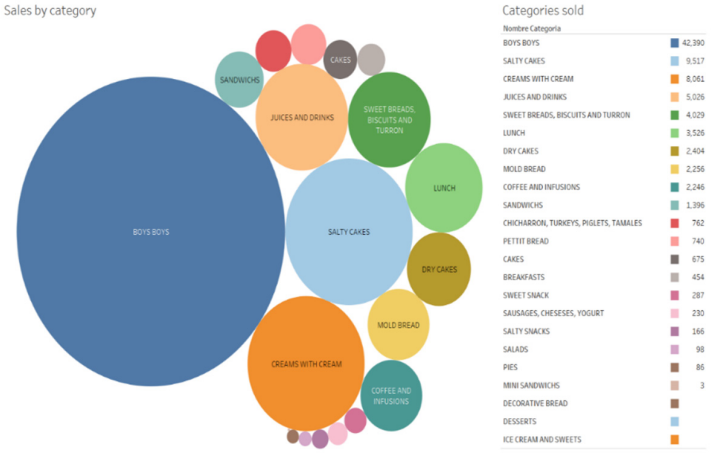


Fig. 8. Graph of total value sold by category in April.

representing 84.17% of total sales by category; in second place was the category of savory cakes with 9,517 units, representing 11.31% of total sales by category; and in third place, the category of cakes with cream with 8,061 units, representing 9.6% of total sales by category.

You want to determine the number of products sold in a month.

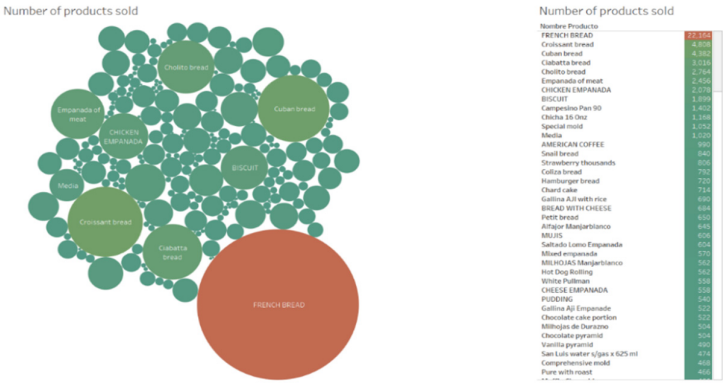


Fig. 9. Graph of products sold in the month of April.

Figure 9 shows the total value sold in April, the most sold product being French bread, which was in first place with 22,164 units sold; in second place was croissant bread with 4,808 units and in third place was cubanito bread with 4,382 units.

We want to determine the total value a customer buys per month.

Figure 10 shows the total value purchased by a customer in the month of April, the Mall del Sur Commercial Center is the customer that has purchased the most products in our company, ranking first with the amount of 4700 products purchased, in second

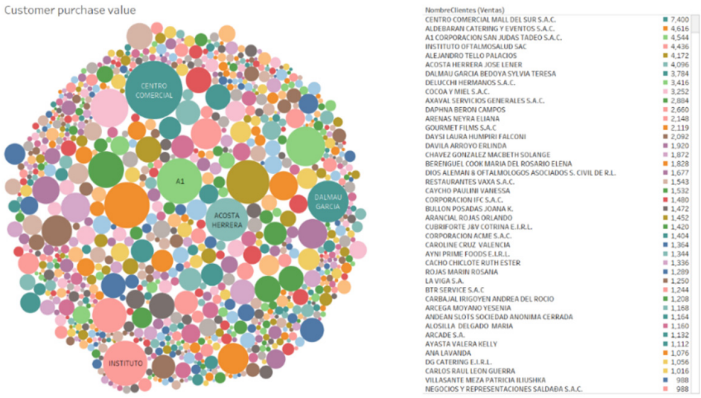


Fig. 10. Graph of total value that a customer buys in the month of April

place was the company Aldebaran Catering and events, with an amount of 4616 products purchased and in third place was the company A1 St. Jude Thaddeus Corporation, with an amount of 4544 products purchased.

2.9 Architecture

Observe in Fig. 11, it can be summarized how the business architecture implemented in Wilton's Bakery, SAC is defined, this architecture is classified in 4 stages. The first stage is the data layer (Data Source), then there is the integration layer (ETL), followed by the Processing and Analysis layer (OLAP), and finally there is the Reports Visualization layer (DashBoards), in this last stage the information worked with value for the company is analyzed.

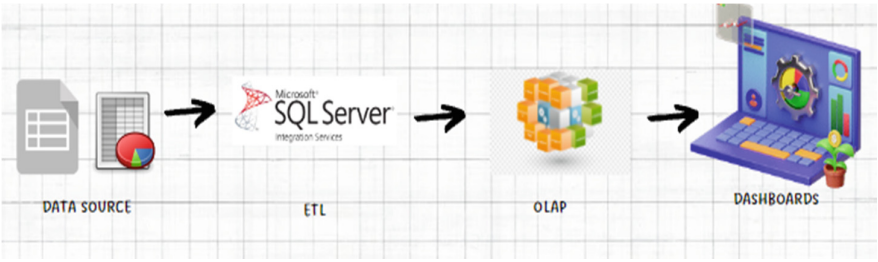


Fig. 11. PDM Architecture Wiltons Baker

3 Results

In the exhaustive analysis of data that were provided by the company, after the implementation and exploitation process of the data cube, the generation of graphical reports was carried out with Tableau software, with this software graphs were created as shown

in Fig. 12 with dynamic data for better detail and scope. Consequently, Tableau allowed to visualize how the sales of the products increased compared to the previous month without the implementation of the Data mart.

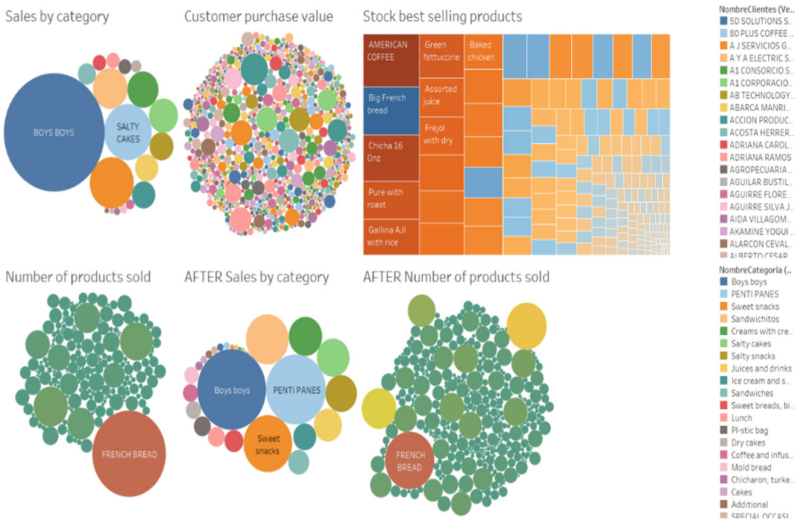


Fig. 12. General Research Dashboard

Figure 13 shows the detailed graphical reports of sales by category, which have increased compared to the previous month, with an estimate of 59,760 in the category of small breads.

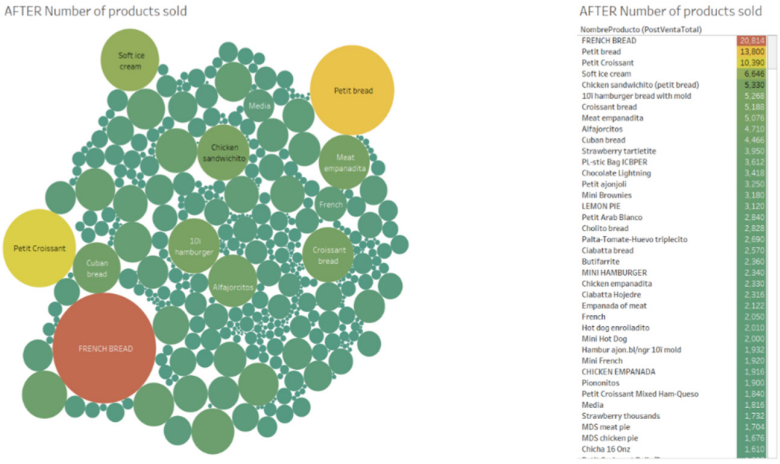


Fig. 13. After Quantity of products sold.

Figure 14 shows the graphic reports in detail of sales by category that have been on the rise compared to the previous month, having a calculation of 59,760 in the category Boys boys.

Fig. 14. After Sales by Research Category

Figure 15 shows in relation to the categories sold, that there are many categories to compare. In the first instance in general, the category of small breads had a sale of 42,390 before the implementation, after that there were 59,760 results, an increase of 30% in that category.

Fig. 15. Comparative categories sold.

According to the number of products sold in Fig. 16, it is highlighted that the PAN CROISSANT product had an increase of more than 45% compared to last month due to the implementation in the company.

Fig. 16. Sales comparisons by product.

4 Conclusions

The design and implementation of the Data mart in the organization with its star model and the application of the Ralph Kimball methodology allowed exploiting the cube using Tableau software, achieving an increase in the sales of the organization of more than 30% in total, within the areas analyzed in the organization. It was found through indicators in the reports, that the company has greater reach with the categories of small breads, since it went from 42 390 units sold to 59 760 units, confirming its best product which is the objective of the company Wilton's SAC linked to its strategic objectives. Croissant bread was also the best-selling product, with a 45% increase in sales.

The sales area at Wilton's Bakery was improved with the design and implementation of the Data Mart through the ETL process and OLAP Cubes, allowing the development of data analysis in real time in a business manner with the use of Tableau Desktop software, generating the graphical report of the results before and after the implementation and design of the datamart. With these results, the decision makers in the decision making area were able to have more information, better and wider scope in their decisions, a better vision about the products to sell and what future strategies can be developed for the benefit of the company.

The customers who purchased the most Wilton's products, with the report as a result of the cube's exploitation according to the category that bought the most were Sd Solutions (Boys boys boys), in second place, Sd Plus Coffe (Pentipnes), in third place, A J Servicios (Sweet snacks), in fourth place, A y A Electric (Sandwichitos) and in fifth place A1 Consorcio S.A. (Creams with cre). It was also found that, in the category of sale of small breads, there was a report of an increase in sales in first place for breads (S/. 59 760), in second place for petit breads (S/. 45 030), in third place for sweet snacks (S/. 29 872) and in fourth place for Sandwichitos (S/. 23 014). Also, we found reports that the company redefined some of its strategic objectives and obtained a post-sale where the boys boys category increased its sales by (S/. 59 760), in second place, pentipanes (S/. 45 030), in third place, sweet snacks (S/. 29 872), in fourth place, sanwichitos (S/. 23 014) and in fifth place, creams with cream (S/. 14 062). In addition, the products that obtained an optimum after sales are French Bread (S/. 20 814), in second place, Petit bread (S/. 13 800), in third place Petit Croisant (S/. 10 390), in fourth place, soft ice cream (S/. 6 646) and petip bread, humburge bread with mokl, Crossant bread and meat bread had an average of S/. 5211. It is concluded that the Data Mart is very efficient in handling large volumes of information, analyzing in real time categories that allow better decision making in organizations.

References

1. Joyanes Aguilar, L.: Big Data, Análisis de grandes volúmenes de datos en organizaciones. Primera Ed. Mexico: Alfaomega Grupo Editor, S.A. de C.V. Pitágoras 1139, Col del Valle, 03100 México D.F. (2013)
2. Acosta Medellín, J., Flórez Lara, D.: Diseño e Implementación de prototipo BI utilizando una herramienta de big data para empresas pymes distribuidoras de tecnología. Universidad Católica de Colombia (2015)
3. Lopez Inga, M., Guerrero Huaranga, R.: Modelo de inteligencia de negocios y analiticaen la nube para pymes del sector retail en Perú. Revista Ingeniería Solidaria **14**, 17 (2018). https://doi.org/10.16925/in.v14i24.2157
4. Santoso, L.W., Yulia: Data warehouse with big data technology for higher education. Procedia Comput. Sci. **124**, 93–99 (2017). https://doi.org/10.1016/J.PROCS.2017.12.134
5. Oumkaltoum, B., Omar, E.B., Aris, O., Loqman, C.: Hybrid e-government framework based on datawarehousing and MAS for data interoperability. Int. J. Adv. Comput. Sci. Appl. **12**(10), 57–64 (2021). https://doi.org/10.14569/IJACSA.2021.0121008
6. Boukobza, A., et al.: Determining the set of items to include in breast operative reports, using clustering algorithms on retrospective data extracted from clinical DataWarehouse. Stud. Health Technol. Inform. **295**, 45–48 (2022). https://doi.org/10.3233/SHTI220656
7. Thivakaran, T.K., Ramesh, M.: Exploratory data analysis and sales forecasting of bigmart dataset using supervised and ANN algorithms. Measur. Sens. **23**, 100388 (2022). https://doi.org/10.1016/j.measen.2022.100388
8. Medina, F., Fariña, F., Castillo, R.: Data mart to obtain indicators of academic productivity in a university | Data mart para obtención de indicadores de productividad académica en una universidad. Ingeniare **26**, 88–101 (2018)
9. Dianderas Alcántara, M.E.: Análisis, Diseño e Implementación de Data Mart de Ventas para optimizar la Toma de Decisiones en una mediana empresa en la ciudad de Lima. Universidad Tecnológica del Perú, vol. 1, p. 113 (2019)
10. Antunes, A.L., Cardoso, E., Barateiro, J.: Incorporation of ontologies in data warehouse/business intelligence systems - a systematic literature review. Int. J. Inf. Manag. Data Insights **2**(2), 100131 (2022). https://doi.org/10.1016/j.jjimei.2022.100131
11. Hoelscher, J., Mortimer, A.: Using Tableau to visualize data and drive decision-making. J. Account. Educ. **44**, 49–59 (2018). https://doi.org/10.1016/j.jaccedu.2018.05.002
12. El-Sappagh, S.H.A., Hendawi, A.M.A., El Bastawissy, A.H.: A proposed model for data warehouse ETL processes. J. King Saud Univ. Comput. Inf. Sci. **23**(2), 91–104 (2011). https://doi.org/10.1016/j.jksuci.2011.05.005
13. Zhou, H., Chen, K., Wang, S.: Two-period pricing and inventory decisions of perishable products with partial lost sales. Eur. J. Oper. Res. (2023). https://doi.org/10.1016/j.ejor.2023.03.010
14. Silva Peñafiel, G.E., Zapata Yánez, V.M., Morales Guamán, K.P., Toaquiza Padilla, L.M.: Análisis de metodologías para desarrollar Data Warehouse aplicado a la toma de decisions. Ciencia Digital **3**(3.4), 397–418 (2019). https://doi.org/10.33262/cienciadigital.v3i3.4.922
15. Sanchez Galvez, F.M.C., Mendoza-Moheno, J.: Alineación del diseño organizacional a una empresa por medio de la aplicación del Modelo Estrella. Boletín Científico INVESTIGIUM de la Escuela Superior de Tizayuca, vol. 8, no. 15, pp. 36–42 (2022). https://doi.org/10.29057/est.v8i15.8751
16. ESheta, O., Nour Eldeen, A.: The technology of using a data warehouse to support decision-making in health care. Int. J. Database Manag. Syst. (IJDMS) **5**(3) (2013). https://doi.org/10.5121/ijdms.2013.5305

17. Zekri, M., Zahaf, S., Gargouri, F.: Specification of the data warehouse for the decision-making dimension of the bid process information system. Procedia Comput. Sci. **159**, 1190–1197 (2019). https://doi.org/10.1016/J.PROCS.2019.09.288
18. Ilha, L., Junior, F.D.J.M.: A construção de um data warehouse utilizando os indicadores educacionais do INEP, p. 31 (2021). https://doi.org/10.1590/SCIELOPREPRINTS.2166
19. Forero Castañeda, D.A., Sanchez Garcia, J.A.: Introducción a la inteligencia de negocios basada en la metodología Kimball, vol. 9, no. 1, pp. 5–17 (2021)
20. Chura, P.C., Yanavilca, A.V., Soria, J.J., Castillo, S.V.: Datamart of business intelligence for the sales area of a peruvian tourism company. In: Silhavy, R., Silhavy, P., Prokopova, Z. (eds.) CoMeSySo 2022. LNNS, vol. 597. pp. 415–429. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-21438-7_33
21. Zadívar, A.R.: Implementación De Un Data Mart Como Solución De Inteligencia De Negocios, Bajo La Metodología De Ralph Kimball Para Optimizar La Toma De Decisiones En El Departamento De Finanzas De La Contraloría General De La República. Universidad De San Martin de Porres, p. 158 (2015)

Intelligent Endoscopic Examination of Internal Openings for Drilling Quality Control

Anton Ivaschenko[1](✉), Vladimir Avsievich[2], Vera Turkova[2], Andrey Belikov[2], and Natalia Chertykovtseva[1]

[1] Samara State Medical University, Chapaevskaya 89, Samara, Russia
anton.ivashenko@gmail.com
[2] SEC "Open Code", Yarmarochnaya 55, Samara, Russia

Abstract. The paper presents an intelligent computer vision complex for drilling quality control using an endoscopic video camera to examine the internal openings and artificial neural networks to identify holes and their defects. The concept is based on a quality guarantor paradigm, which provides statistical analysis of the trajectories of repeatable actions that characterize manufacturing process. The hardware part of the complex is built on a 5-axis machine, the working area of which allows access to the surfaces of the controlled part. As part of this solution, two new algorithms were developed: an algorithm for precise positioning of the endoscope to collect the necessary initial data, and a computer vision algorithm for detecting defects. For the second algorithm, an array of initial data was specially collected, which made it possible to train and subsequently use the neural network to monitor the inner surface of the holes. The resulting complex is designated to be used in manufacturing as a part of a quality control decision-making support system. #COMESYSO1120.

Keywords: Computer Vision · Quality Guarantor · Quality Control · Artificial Intelligence · Smart Factory · Industry 4.0 · Bioengineering

1 Introduction

Manufacturing production control is an important task in implementing the concept of Industry 4.0. Usually, for such control, especially or products with complex shapes, specialized sensors and devices are used to make the necessary measurements and compare their results with the data of design documentation. However, in the case of small-scale or pilot production, the implementation of such control is difficult due to the high cost of measuring equipment and accessories.

It is promising to use modern computer vision technologies instead, which allow recognition the geometric parameters of products with accuracy sufficient to identify gross flaws and defects. For example, intelligent computer vision can help control the quality of internal machining, such as when forming holes in a complex housing. Identification of the facts of the absence of the necessary internal openings, their incorrect location or shape can facilitate quality control of products in case of their covert placement.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 397–403, 2024.
https://doi.org/10.1007/978-3-031-53552-9_35

However, to implement this kind of technical diagnostics, it is necessary to provide video recording in the holes and cavities of products, which requires solving the complex problem of automated camera immersion. This problem can be solved by analogy with medical solutions, for example, by using an endoscope camera. This article proposes an original solution to this problem, based on the combined use of an endoscope control system and artificial neural networks.

Below there is presented a new solution developed for drilling quality control by integrating an endoscopic camera system with artificial neural networks. After an overview of existing solutions there are proposed the methods of examining the inner surface of the holes using an endoscopic camera and results of their practical implementation, which illustrate the positive effect of a quality guarantor paradigm.

2 State of the Art

According to the concept of modern Industry 4.0 [1–3] intelligent computer vision systems can be widely implemented for manufacturing quality control. In practice, they can be used either an autonomous system, as part of robotic quality gates, or together with the operator to support automated decision-making support. The second solution is preferable, as soon as it provides rational combination of human and artificial intelligence [4, 5].

The use of computer vision technologies for quality control of manufactured products is carried out at the enterprise in accordance with modern requirements and quality management technologies [6, 7]. The visual data obtained is statistically processed and summarized in order to provide an adequate and reliable result. Statistical methods of data processing [8, 9] can reduce the error in the use of computer vision in e.g. low light conditions.

In particular, the task of quality control using a computer vision system is difficult and challenging to control holes [10, 11], especially in the case of their complex shape and location. Existing solutions can help detecting reference holes accurately and robustly. In addition to it the system can identify the characteristics of hole quality that require instrumental control including the hole size, circularity or roundness error, burr formation, surface roughness, etc.

This problem is especially specific for drilling of dissimilar material combinations [12] having different mechanical properties that might cause defects of holes. Identification of these defects is an essential problem of drilling quality management and control. Considering the specifics of used materials and implemented technologies this problem is challenging for aircraft industry [13, 14].

On the software level modern computer vision systems implement artificial neural networks for production quality control [15, 16]. The use of this toolkit requires an individual approach to solve each specific task. For example, to analyze internal holes, the neural network needs to be additionally completed and retrained.

To solve the problem of internal openings for manufacturing quality control, we can offer an analogy with biological objects and the healthcare industry [17], for example, by using a robotic endoscope camera [18].

In general, implementation of computer optics for drilling quality control is known in computer science [19, 20], but however, there is no universal solution based on artificial

intelligence. This paper proposes such a solution implemented for machine-building production.

3 Methods

To build a quality control system for the drilling production process, an automated complex for monitoring production parameters was developed that implements the ideas of a quality guarantor. The main solution is based on the implementation of an endoscopic camera, which allows examining the inner surface of the holes. The image obtained by the endoscope is input to a computer vision subsystem that uses an artificial neural network to identify defects.

As part of this solution, two new algorithms were proposed: an algorithm for precise positioning of the endoscope camera to collect the necessary initial data, and a computer vision algorithm for detecting defects. A special feature of the solution is the ability to use the drill movement track, specified by the technological process, to control the movement of the monitoring endoscope. For the second algorithm, an array of initial data was specially collected, which made it possible to train and subsequently use the neural network to monitor the inner surface of the holes.

The hardware part of the complex (see Fig. 1) is built on a 5-axis machine, the working area of which allows access to all surfaces of the controlled part, except for the base.

Fig. 1. Quality control complex with implemented endoscope (at the left) and the video results obtained by a computer vision system (at the right).

The approach to the base of the part will be provided by turning it over manually. Accurate positioning of the part on the machine's coordinate table and its correct turning by the user will be ensured by specialized equipment that makes it impossible to incorrectly position the part, as well as by using a neural network that recognizes the correct position of the part before starting to control a specific list of holes.

The main measuring equipment of the quality control complex is a video endoscope, fixed on the moving part of the machine. The device is connected to the built-in personal computer via UART port, and is controlled by the G-code protocol.

The quality monitoring algorithm implemented by a complex is based on the control program that was used to manufacture the part of the machine, so we get an absolutely accurate trajectory for the endoscope. The list of holes and their intersections, as well as parameters for measurement, is known in advance. The check is carried out sequentially, in accordance with the specified order.

The measurement accuracy is determined mainly by the camera resolution and depends on the results of training the neural network. To implement an effective quality control system based on machine vision, it is necessary to work out in detail the list of controlled parameters and compile corresponding examples of defects in the form of images as part of the training dataset. Note that since it is difficult to accumulate enough images illustrating defects, the missing images can be obtained by simulation modeling.

4 Results

The first way to detect holes is to use the SimpleBlobDetector algorithm. SimpleBlobDetector is a computer vision algorithm used to detect simple objects in images. This detector works by looking for prominent areas in the image, which can be considered as "spots". To filter out unwanted objects, various criteria can be applied, such as minimum and maximum size, shape, or intensity.

After preliminary transformation and the application of the SimpleBlobDetector algorithm, it was possible to detect darkened elements in the image, which were holes or parts of holes. Due to the presence of glare on the edges of the holes, the detected objects are separated by this glare. In this regard, the detected objects that overlap each other are combined into one larger object in order to correctly identify the entire hole (see Fig. 2).

To determine the direction of the side hole (when crossing the sub holes of the manufactory unit with a complex shape), it is possible to resort to the PCA method. PCA allows finding the principal components in a multidimensional dataset. Eigenvectors or principal components are sorted based on their respective eigenvalues in descending order. This allows you to select the most important principal components that cover the most significant variance in the data. Thanks to this approach, it was possible to determine the presumed direction of the hole. However, this approach proved to be too unstable, as it is highly dependent on changes in illumination in the inner hole.

For this reason, the second option for detecting holes was considered, and, subsequently, was chosen as the main one. The second way to detect holes is a neural network approach based on the use of the YOLOv4 neural network (see Fig. 3).

The collected dataset involved 2213 images, among which 2166 images were used for training. As a result of training, the result of the mAP metric was 89%, which is acceptable for implementation in practice as a part of an intelligent decision-making support system for drilling quality control.

Currently, the system allows you to work with aluminum products under normal workplace lighting. When implementing the solution in a real application, the results

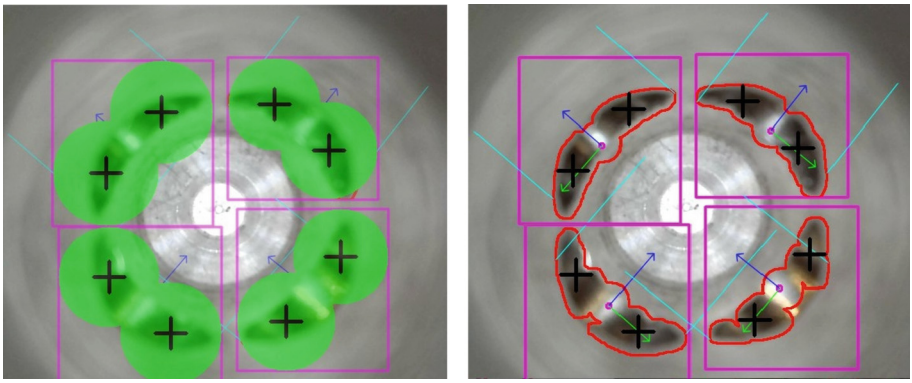


Fig. 2. The result of the detection of parts of the holes and their combination.

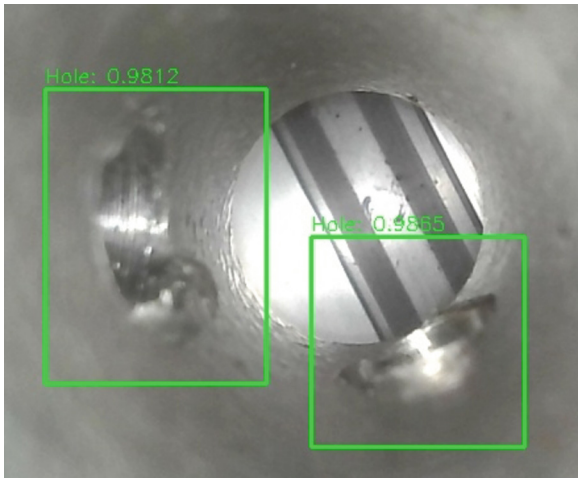


Fig. 3. Holes' identification result.

of computer vision are duplicated on the screen of the dispatcher, which allows the system to be implemented in the format of intelligent decision support. The actions of the dispatcher are also recorded and can be used in the future for additional training or retraining of the neural network.

5 Discussion

The neural network approach to identify and control holes in the process of drilling quality control is characterized by the following benefits:

- accuracy and performance: YOLOv4 provides high accuracy while maintaining real-time or near-real-time inference speed, uses efficient network architecture and optimization techniques, including multi-scale training and testing, to achieve high inference speed;

- flexibility: in terms of model configuration and deployment options it supports various input data sizes, allowing you to find a compromise between accuracy and speed;
- open source: YOLOv4 is open source, making its codebase available for customization and integration into projects.

Limitations of the proposed approach are concerned with the low accuracy of computer vision methods. The use of an endoscope allows you to identify only visually observable defects, while it is better to use more accurate methods and measuring instruments to identify geometric deviations, poor surface quality, roughness and other controlled parameters.

The scope of application of the endoscope is limited by its dimensions, but for small products they are sufficient. In addition, in a quality control system based on computer vision, quite a lot depends on the results of training the neural network, the volume and variety of the training dataset and algorithms for interpreting the results obtained by artificial intelligence.

6 Conclusion

Research results prove the possibility to implement an intelligent computer vision system for production quality control of internal openings with complex structure. Next steps are considered with an extension of endoscopic examination use in manufacturing quality control. Future efforts are aimed at improving the efficiency of endoscope motion algorithms to reduce quality control time, as well as improving the accuracy of defect detection for different types of abnormalities.

References

1. Roblek, V., Meško, M., Krapež, A.: A complex view of Industry 4.0, SAGE Open, vol. 6, pp. 1–11 (2016)
2. Mamad, M.: Challenges and benefits of industry 4.0: an overview. Int. J. Supply Oper. Manag. **5**, 256–265 (2018)
3. Kumar, A., Kumar, S.: Industry 4.0: evolution, opportunities and challenges. Int. J. Res. Bus. Stud. **5**(1), 139–148 (2020)
4. Ivaschenko, A., Golovnin, O., Avsievich, V., Sitnikov, P.: Intelligent quality guarantor model for computer vision based quality control. In: Goar, V., Kuri, M., Kumar, R., Senjyu, T. (eds.) Advances in Information Communication Technology and Computing. LNNS, vol. 628, pp. 1–9. Springer, Singapore (2023). https://doi.org/10.1007/978-981-19-9888-1_1
5. Surnin, O., et al.: Augmented reality implementation for comfortable adaptation of disabled personnel to the production workplace. In: Proceedings of the 35th Annual European Simulation and Modelling Conference, pp. 64–69 (2021)
6. Blanco-Encomienda, F., Rosillo-Díaz, E., Muñoz-Rosas, J.: Importance of quality control implementation in the production process of a company. Eur. J. Econ. Bus. Stud. **4**, 240–244 (2018)
7. Dominguez, O.: Quality management: the heart of the quality assurance/quality control process. Spectrosc. Eur. **33**(1), 27–32 (2021)
8. Fisher, N., Tanaka, Y., Woodall, W.: The road to quality control: the industrial application of statistical quality control by Homer M. Sarasohn, Wiley, 160 p. (2019)

9. Moru, D., Borro, D.: A machine vision algorithm for quality control inspection of gears. Int. J. Adv. Manufact. Technol. **106** (2020). https://doi.org/10.1007/s00170-019-04426-2
10. Aamir, M., Giasin, K., Tolouei-Rad, M., Vafadar, A.: A review: drilling performance and hole quality of aluminium alloys for aerospace applications. J. Market. Res. **9**, 12484–12500 (2020)
11. Mei, B., Zhu, W., Yan, G., Ke, Y.: A new elliptic contour extraction method for reference hole detection in robotic drilling. Pattern Anal. Appl. **18**, 695–712 (2015)
12. Barik, T., Sarangi, S., Pal, K.: Assessment on hole quality during drilling of Al/CFRP stack. In: Advances in Unconventional Machining and Composites, pp. 757–770 (2020)
13. Kawano, F., Toledo, C., Barbosa, G., Sagawa, J., Shiki, S.B.: Improving the quality assessment of drilled holes in aircraft structures (2023). https://doi.org/10.21203/rs.3.rs-2705166/v1
14. Teti, R., Segreto, T., Caggiano, A., Nele, L.: Smart multi-sensor monitoring in drilling of CFRP/CFRP composite material stacks for aerospace assembly applications. Appl. Sci. **10**(3) (2020). https://doi.org/10.3390/app10030758
15. Jiang, L., et al.: MA-YOLO: a method for detecting surface defects of aluminum profiles with attention guidance. IEEE Access 1 (2023)
16. Kottner, L., Mehnen, J., Romanenko, D., Bender, S., Hintze, W.: Process monitoring using machine learning for semi-automatic drilling of rivet holes in the aerospace industry. In: Behrens, B.A., Brosius, A., Hintze, W., Ihlenfeldt, S., Wulfsberg, J.P. (eds.) Production at the Leading Edge of Technology, pp. 497–507. Springer, Heidelberg (2021). https://doi.org/10.1007/978-3-662-62138-7_50
17. Ang, E.C., Suandi, S.A.: Smart manufacturing with an artificial neural network to predict manufacturing healthiness. In: 2019 IEEE 15th International Colloquium on Signal Processing Its Applications (CSPA), pp. 120–123 (2019)
18. Boese, A., et al.: Endoscopic imaging technology today. Diagnostics **12**, 1262 (2022)
19. Sitek, P., Katunin, A., Kajzer, A.: Analysis of drilling process of composite structures – Part II: quality control of drilled holes. Modelowanie Inżynierskie **24**, 95–105 (2015)
20. Singh, M., Dhiman, S., Singh, H., Berndt, C. Assessment of positional error and hole quality during vibration-based drilling of aerospace alloy. J. Mech. Sci. Technol. **35**, 1–10 (2021)

Foreign Languages Teaching Technologies and Methods in Traditional and Distance Forms of Learning

E. D. Lavrinenko, L. A. Bayrak, J. I. Erkenova(✉), I. Sh. Kappusheva, O. A. Frolova, and N. A. Bystrov

MIREA-Russian Technology University, Moscow, Russia
erkenova@mirea.ru

Abstract. The article is devoted to the use of various technologies for teaching foreign languages. An overview of certain methods used for teaching foreign languages is given. The article describes some technologies and their advantages and prospects. Particular attention is paid to the analysis of different forms of teaching in the form of research activities, for example to a survey.

Keywords: Teaching Technologies · Distance Learning · Project-Based Learning · Language Portfolio · Quality of Teaching

1 Introduction

The concept of learning technology is directly related to a large number of very different methods and techniques, with the help of which new knowledge can be obtained. The concept of "technology" comes from the Greek word "techne", which means "art, skill" and "logos", which means "teaching, knowledge. "Based on the proposed interpretation of the concept, it can be assumed that in the general sense, "technology" means scientific knowledge, mastery skills. If you turn to the dictionary of the Russian language, you can find that the word "technology" means a set of techniques used in any business, skill, art. As for the topic that is considered, there are various methods of teaching foreign languages, which are discussed further in this article. Such technologies as the project method, collaborative learning, the use of a language portfolio, audiovisual and computer technologies are usually referred to as learning technologies [1]. All these technologies can be used in both traditional and distance learning. Moreover, computer and information and communication technologies are especially widely used in distance learning.

2 Methods

The existing methods are considered in more detail in this section.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 404–410, 2024.
https://doi.org/10.1007/978-3-031-53552-9_36

2.1 Project-Based Method

The project-based method is based on the development of students' cognitive skills, the ability to independently construct their knowledge, the ability to navigate in the information space, the development of critical and creative thinking [2]. This is a set of educational and cognitive techniques that allow solving a particular problem as a result of independent actions of students with the obligatory presentation of these results. The project-based method is aimed at developing the cognitive skills and abilities of students, as well as their ability to create their own knowledge and navigate the information space. The main feature of this approach is the use of educational and cognitive techniques that provide an opportunity to solve various problems through the independent actions of students and the presentation of results. Project-based technology includes a variety of methods of research and search for solutions, as well as creative approaches. The advantage of the project-based method is, first of all, that the time limitation of the lesson can be overcome by the extracurricular independent work of students. The project-based method is always focused on the independent activity of students performing individual, pair or group work. This approach is seamlessly combined with group methods. Problem solving is the main goal of the project-based method. This goal is achieved through the use of various methods and training resources, as well as through the integration of knowledge and skills from different areas of scientific, technical, technological and creative knowledge. The results of project implementation should always be "tangible", i.e. concrete and ready to use.

2.2 Collaborative Learning

This training in small groups allows teachers to pay more attention to each student, evaluate students' understanding of the material and help solve problems that have arisen. Also, this approach helps to develop students' social skills and abilities, such as communication, cooperation, leadership and self-confidence [3]. There are many methods and strategies for organizing collaborative learning, such as round-robin discussion, brainstorming, role-playing, feedback, co-writing, etc. Each of these methods encourages students to actively participate in learning, jointly seek solutions, analyze and evaluate the results. Collaborative learning is currently becoming more popular and in demand in modern education, as it allows students to better prepare for the modern labor market, where it is required to work in a team, share knowledge and ideas, and solve complex problems. However, it should be understood that the above mentioned methods and strategies are difficult to implement or not implemented at all with a completely distance form of education, with the absence of live direct interaction with other participants. Such methods and strategies can be easily implemented with traditional forms of education or with mixed education combining on-line and off-line activities.

2.3 The Use of the Language Portfolio

In modern methodology, a language portfolio (from French porter - to wear + feuille - leaf) is defined as "a set of working materials that represent the student's learning experience in mastering a foreign language [4]. Such a set of materials gives students

and the teacher the opportunity to independently or jointly analyze and evaluate the amount of work and the achievements in language and culture learning, the dynamics of mastering the language being studied.

The goals of this technique are:

- to evaluate the student's abilities, skills, personal experience with self-assessment, mutual assessment and assessment by adults;
- to explore your own development over a period of time;
- to set realistic learning goals for the near future;
- to develop the skill of self-reflection in students.

The objectives of this technique are:

- to note language abilities and skills;
- to investigate the development of the student over time;
- to study the features of oneself as a person and develop self-confidence;
- to develop writing skills.

2.4 Audiovisual and Computer Technologies

Computer training is based on the ideas of programmed learning and today has a significant impact on all aspects of the educational process due to the mass computerization of secondary and higher schools, the creation of computer programs for academic disciplines, including learning foreign languages, using the capabilities of the Internet in the classroom and during the independent work of students. In Russian linguodidactics, the term "Computer linguodidactics", proposed in 1991 by K.R. Piotrovskaya and developed in relation to teaching languages [5] is used to specify scope of questions related to the use of computer technologies in teaching process. At present, the following types of computer training programs have become widespread: those devoted to mastering individual sections of the language system; devoted to the formation of speech skills; programs evaluating the level of language proficiency; sociocultural programs that introduce the culture of the country of the language being studied. The use of Internet technologies allows the teacher to model alternative learning strategies that take into account the psychological characteristics of students. Only a computer makes it possible to effectively combine auditory, visual and kinesthetic learning styles. The Internet has opened up truly unlimited possibilities for imbedding foreign language teaching with modern authentic content. The latest issues of world-famous newspapers, videos with the latest news, correspondence, online discussions with native speakers - ten years ago, one could only dream of this [6].

2.5 Distance Learning

Separately, there is distance learning, which is more a form of education than a technology. This form of training may include various methods and technologies, including those mentioned above. Not so long ago, the remote format was considered complex and incomprehensible. But today it is actively used in all educational areas - for example, in the system of school and higher education, in the development of companies' personnel and in online schools [7].

By online learning, many still understand the work of a teacher and a student at a distance. In fact, this concept is much broader. Distance learning technologies are a combination of the latest information methods and forms of development that provide the educational process at a distance. Distance no longer prevents live communication, interaction and practical experience. Distance learning is learning at a distance. The distance learning system allows you to acquire the necessary skills and new knowledge using a personal computer and access to the Internet. The location of the computer does not matter, so you can study at home, at work, in the on-line class of one of the distance learning centers, as well as in any other place where there is a PC with an Internet connection. This is the most important advantage of distance learning over traditional forms of learning.

2.6 Conclusions

Thus, modern technologies for teaching foreign languages allow us to conclude that at present there are two main ways to intensify the learning process: the one - via the maximum use of technical means, the other - via the activation of the reserves of the personality of each student.

It can be asserted that the development of intensive methods, which provide for the seamless inclusion of technical means in the training system, will be the most promising direction for improving the methodology in the coming years in the direction of intensifying the educational process [8]. There is no doubt that the most promising in the coming years will be learning technologies that provide for the use of computers and various forms of distance learning in the educational process, since the world does not stand still and improvements and discoveries related specifically to distance and computer technologies play a leading role. It is worth noting that, in addition to progress in this area, there are other advantages of this type of training [9]. It is practical and convenient to study from anywhere in the world.

3 Results

In this study, the use of various technologies in teaching foreign languages is considered on the basis of the Russian Technological University (RTU MIREA).

For the implementation of experiment, 60 students of the 2nd year of the Institute of Radio Electronics and Informatics were involved and interviewed.

The survey consisted of rather important questions that allowed us to consider the full picture. It is clearly seen what ideas about the various methods related to the technologies of teaching foreign languages people have (Figs. 1 and 2).

During the ascertaining stage of the experiment, we found that the majority of respondents did not particularly think about the new technologies and were accustomed to classical forms of education.

It should be noted that the majority did not particularly attach importance to this issue (it can be seen from the results of the survey). In fact, this is quite normal, since many people learn without thinking about how to do it. They decide whether this format suits them or not during the process of learning. In most cases, especially when it comes to an

Fig. 1. A visual representation of students' awareness about various technologies for teaching foreign languages.

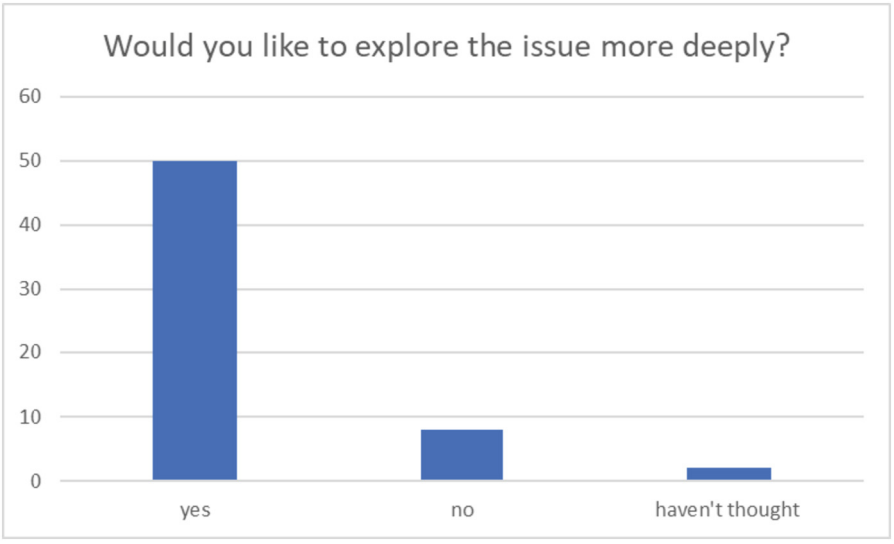


Fig. 2. Interest of respondents in this topic.

educational institution, there is no choice, since there is a certain academic programme, which, as a rule, does not deviate from.

The survey did not end at this stage. Additional question was asked, which directly relates to the personal desire of students to understand this topic.

4 Discussion

In the modern world, technological progress has affected all spheres of life, including education. Today, the learning of foreign languages has become much more accessible thanks to new learning technologies. However, do not forget about the traditional methods of learning languages, such as classes with teachers in the classroom and reading books in a foreign language. After all, these methods provide an opportunity for in-depth acquaintance with the language and its grammar, as well as teach new words and expressions that are not always found in online dictionaries. Technology makes it possible to speed up the learning process, for example, through online courses or mobile applications. It is important to choose acknowledged resources that really help in learning the language and correspond to the level of difficulty.

Modern technologies also make it possible to communicate with native speakers via the Internet. This is a chance to speak the language and improve your pronunciation and intonation, which is especially important for speaking practice. Also, online communication with native speakers contributes to broadening one's horizons and learning about the culture and traditions of other countries. But no less important is the independent study of foreign languages. The use of technology for independent learning includes listening to podcasts, reading articles in foreign languages, and watching films in their original language. Thus, technologies for teaching foreign languages have many advantages, but they are not a panacea. All teaching methods can be used in combination with each other for more effective learning of foreign languages. It is important to choose acknowledged resources and follow the learning process in order to improve your knowledge.

5 Conclusion

It can be concluded that there are a huge variety of different technologies that assist in achieving success in learning foreign languages. In the course of linguistic training, these methods develop not only some communication skills, but also affect a person from the point of view of personality. Such qualities as independence, self-organization and self-control arise. In conclusion, we can say that foreign language teaching technologies simplify the learning process and make it more accessible. Today, learning languages has become much easier, due to a huge number of resources and programs for learning languages. It is important to understand that technology cannot completely replace classical teaching methods like reading books and lessens with teachers. However, they can and should complement the learning process, making it more interactive and fun. Today, specialized applications are available that allow you to learn languages in the form of a game, social networks for communicating with native speakers, video courses and other learning formats that make it possible to choose the most convenient way for everyone. Researches also prove that the use of technology in the learning process can be safely considered one of the key factors that ensure successful language learning. After all, modern learning and teaching technologies allow you to acquire new knowledge much faster and better retain it in memory. However, it should be noted that there are some disadvantages in the use of technology in teaching foreign languages. For example, some programs may be incorrect and in such a case may lead to improper acquisition of

language skills. It is also obvious that if we use programs in a language that we do not know well, this can also become an obstacle in the learning process.

Thus, technology is an integral part of modern language learning. They not only improve the quality of learning, but also make it more accessible and interesting. This is an important contribution to learning that is becoming more and more significant every year.

References

1. Project Method. http://lyceum144.ru/media/2014/03/metodologicheskie-osnovyi-proektnoy-deyatelnosti.pdf. Accessed 9 Sept 2020
2. Shokanova, R.D., Tarasova, E.N.: The reproductive method in teaching foreign students the Russian language and its innovative aspects. Russian Technol. J. **9**(3), 98–107 (2021)
3. Sivarajah, R.T., Lee, J.T., Richardson, M.L.: A review of innovative teaching methods. Acad. Radiol. **26**, 101–113 (2019)
4. Arkusova, I.V.: Modern pedagogical technologies in teaching a foreign language (structural-logical tables and application practice). Moscow Social Pedagogical Institute (MSPI), Moscow, Russia (2014)
5. Technology "Language Portfolio" and its use at the lessons foreign language in grades 6–7. http://gymnasium.pruzhany.by/wp-content/uploads/2020/06/Технология-ЯП-и-ее-использование-на-уроках-ИЯ.pdf. Accessed 9 Sept 2020
6. Shchukin, A.N.: Modern intensive methods and technologies for teaching foreign languages. Filomatis, Moscow, Russia (2014)
7. Pikhart, M., Al-Obaydi, L.H.: Potential pitfalls of online foreign language teaching from the perspective of the university teachers. Heliyon **9**, 45–51 (2023)
8. Kondrateva, I.G., Valeev, A.A.: Technology of Students' professional competences formation while teaching a foreign language. Procedia. Soc. Behav. Sci. **191**, 2250–2253 (2015)
9. It is impossible to teach you English. https://ik-ptz.ru/en/dictations-on-the-russian-language--class-3/vas-nevozmozhno-nauchit-angliiskomu-yazyku-zamyatkin-zamyatkin-n-vas.html. Accessed 9 Sept 2020

The Usage of Distance Educational Technologies for Learning English

V. A. Sogrina, D. A. Stanovova, J. I. Erkenova(✉), I. Sh. Kappusheva, F. A. Nanay, and A. V. Filipskaya

MIREA-Russian Technology University, Moscow, Russia
erkenova@mirea.ru

Abstract. The article considers the use of distance learning technologies in the course of foreign language training at the Russian Technological University (RTU MIREA). As part of the ongoing experimental research, the authors evaluated the possibilities of improving communication skills in intercultural interaction, contributing to the solution of professional problems.

Keywords: Foreign Language Training · Distance Learning · Video Materials · Communication Skills

1 Introduction

In the modern world, digitalization has affected all spheres of human professional activity: economic, political and, first of all, educational one [1]. Controversy on the world scene, as well as the difficult epidemiological situation associated with the coronavirus pandemic (COVID-19) and self-isolation of citizens, have challenged many areas of public life, forcing society to adapt to change. The former traditional education system had to be adapted to the new conditions as soon as possible. Not surprisingly, distance learning is one of the most relevant teaching methods today. It includes, among other things, the use of electronic media for learning and information exchange. The idea of using electronic resources and the Internet in the educational process appeared long before the pandemic, since the role of Internet technologies in education [2] has seriously increased with the onset of the 21st century due to the continuous and rapid development of Internet technologies and information and communication technologies (hereinafter referred to as ICT).

In our country, information educational technologies have received intensive development and have begun to confidently gain their place in the educational process along with traditional forms of education, so we are increasingly talking about distance learning. With the development of technology and progress in the field of online learning, people use distance education, primarily to learn foreign languages, such as English more often. Numerous studies demonstrate that the introduction of new Internet technologies into the education system increases the efficiency of the educational process, makes foreign language lessons more productive and effective [3].

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 411–417, 2024.
https://doi.org/10.1007/978-3-031-53552-9_37

Currently, there is still no single definition of the term distance learning among theorists and practitioners. However, the Federal Law “About Education in the Russian Federation” gives the following definition of distance education: “Distance educational technologies are understood as educational technologies implemented mainly with the use of information and telecommunication networks with indirect (at a distance) interaction between students and teachers” [4]. That means that distance learning assumes that training and control over the mastering the materials and knowledge take place via the Internet computer network, using on-line and off-line technologies. Modern information technology tools make it possible to use various forms of material presentation in teaching: verbal and figurative (graphics, sound, animation, video). Computer-based training and monitoring programs assist students in mastering the educational material faster and deeper.

In this article, we are discussing the experience of using distance learning technologies for learning English, the advantages and disadvantages of this method, as well as the best resources that can be used to learn the language.

2 Methods

Let’s take a closer look at some of the common benefits of distance learning.

1. Flexibility. Here we mean that thanks to distance learning, a student can study at a convenient time, in a convenient place and at a convenient speed for completing tasks.
2. Parallel learning. By choosing a distance learning form, students can decide for themselves the conditions and pace of learning, which allows them to work simultaneously with learning.
3. Long-range action. This property means that the distance to the educational institution is not an obstacle. One can learn from any place of the world.
4. "Asynchrony”. This advantage implies a convenient work schedule for each of the parties.
5. Coverage or mass character.
6. Economy and practicality. Distance education is much cheaper than traditional full-time education.
7. Such a parameter as internationality indicates the possibility of exporting and importing educational services.
8. Modularity of distance education. The principle of modularity allows the formation of a complete curriculum, meeting individual or group needs, from individual educational courses.
9. Specialized quality control of certification training in the format of standardized testing.
10. There are benefits in the field of sociality as well. This form of education relieves social tension. Regardless of place of residence, income level and other criteria, distance learning makes it possible for anyone to get an education.
11. An educator who acts as a tutor, not a censor.

All these advantages are perfectly applicable to the study of foreign languages, namely English. When discussing the advantages of distance form of learning in learning English, all students note the flexibility and accessibility of this method. Students can study anywhere and at any time, choosing a convenient class schedule, which allows them to adapt to their personal schedule and daily routine, not requiring a physical presence in educational institutions. There are a huge number of courses and materials for learning English on the Internet, which allows to choose the most suitable one for each student. A lot of materials for distance learning of English are currently online. This means that one can always find new and relevant content. The use of distance technologies for learning English saves not only time, but also money. Learning English online allows students to learn materials at their own pace and according to their own level, and can also provide students with a unique opportunity to practice their communication skills in real time.

But despite the advantages, distance learning has some disadvantages as well.

1. And oddly enough, one of the advantages of distance learning turns out to be its disadvantage for some students. Earlier we have mentioned such benefits of distance learning as distance control and supervision. Besides we have mentioned that a teacher is more of a tutor than classical teacher in distance learning. But for some students this limited contact with the teacher is one of the main drawbacks of distance learning. This can lead to a lack of practical application of knowledge.
2. Distance learning requires self-sufficiency and discipline from students.
3. In addition, a computer or laptop is required, as well as access to the Internet.
4. Online learning does not provide the level of social interaction that classroom learning can provide. Distance learning is not suitable for developing communication skills.
5. The problem of user authentication when checking knowledge, which means the impossibility of 100% control over students' knowledge.

In Russia, distance learning at the moment will not be able to completely replace traditional forms of education. In any case, the use of fully distance learning for obtaining the first higher education and a diploma of higher education seems impossible so far. In our country, distance learning is more likely to be used for self-education, additional courses, advanced training, and also as an auxiliary tool for the traditional form of education. Various online language learning tools are a prime example of this.

Now let's have a look at various online resources where you can learn English remotely. The resources below are the best and most popular in the field.

There are a lot of free and paid mobile apps for learning English. They can offer different levels of difficulty and learning, ranging from beginner to advanced. One of the most downloaded apps for learning English and other languages is "Duolingo", that allows for example to learn a language via tests where you have to translate familiar words. "Babbel" is a language learning app that offers a large database of words and phrases that you can use in learning language.

Online courses are one of the most popular methods for distance learning of English. They often offer language proficiency tests to determine a student's level, allowing students to learn at their own pace and according to their own level. "EF English Live" allows to learn English in online classes where you can get real-time feedback and communicate with your colleagues using your computer. "Rosetta Stone" is a paid online course that offers quality English teaching and provides lessons with native speakers.

There are many free websites that offer online resources for learning English. They include online dictionaries, tests, video tutorials, games, and reading and listening materials. The most popular site is "BBC Learning English", which provides a huge amount of free English learning material for all age groups. Another equally viewed site is "English Central" – an online English course that offers numerous videos on a variety of topics.

Further in this article, we are considering the elements of distance learning as an auxiliary tool for traditional full-time education in an institution of higher education - MIREA – RUSSIAN TECHNOLOGICAL UNIVERSITY (RTU MIREA).

To study the effectiveness of distance learning, we have conducted special research. We interviewed 50 s year students of the Institute of Radio Electronics and Informatics and the Institute of Artificial Intelligence. During the ascertaining stage of the experiment, we found that the majority of respondents did not know about the possibility of distance learning using YouTube channels (Fig. 1).

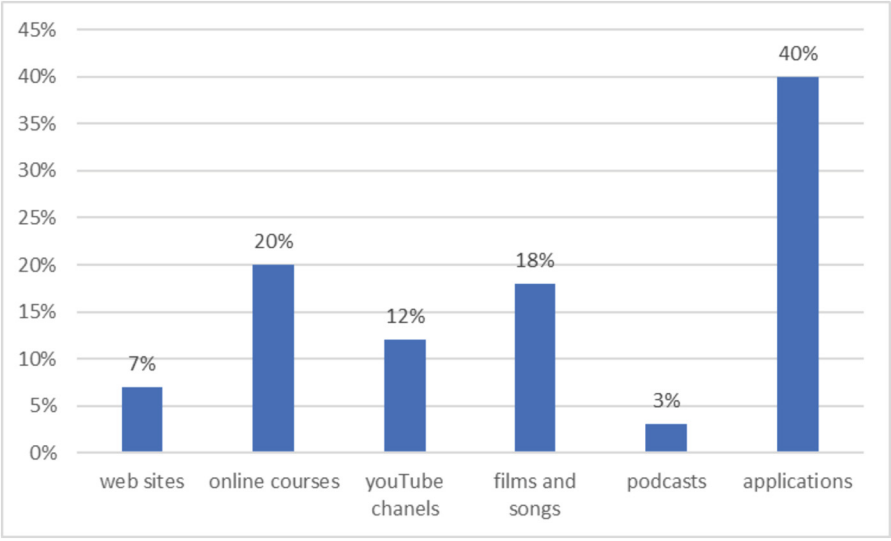


Fig. 1. The use of distance technologies for learning English.

At RTU MIREA, students study a foreign language for the first two years of their study - in the first and second years. During the first year of training students are studying so called general English. As for the second year – students are studying technical English according to their specialization. As part of the professional foreign language training of students of the Institute of Radio Electronics and Informatics in RTU MIREA one of the topics of the fourth semester is "Transistors". The materials of the unit about transistors consist of authentic texts, lexical and grammatical tasks and video materials. Using video materials in the process of teaching and learning foreign languages can be considered as one of the distance learning methods. Using the video material "Transistors, how do they work?" as an example, a study was conducted on the efficiency and effectiveness of this distance learning method. The experimental study (formative stage) was carried

out in 3 steps. At each step we empirically evaluated the respondents on the following issues: to hear (simple perception of sounds), to listen (to try to have the main idea from the flow of foreign language information), to analyze (to make any notes about watched video), to imitate (to try to reproduce what one has heard from memory) and to understand (to fulfill the tasks and be able to retell what one has heard). The survey proved that the division into stages of such an activity as watching video files effectively affects the formation of a student's communication skills [5]. It is worth noting that during the initial viewing video files, the indicators "to hear" and "to understand" vary greatly, thereby demonstrating that the respondents can perceive foreign speech, but cannot immediately formulate the main idea. Re-watching the video material allowed students to take notes to analyze the presented material, thereby increasing the number of students who were able to partially imitate the text they heard. At the final stage of watching the video file, students were proposed to use subtitles. This made it possible to follow the information visually and correlate by ear the material presented in the lesson, bringing the understanding of the speech to almost 100% of the result (Fig. 2).

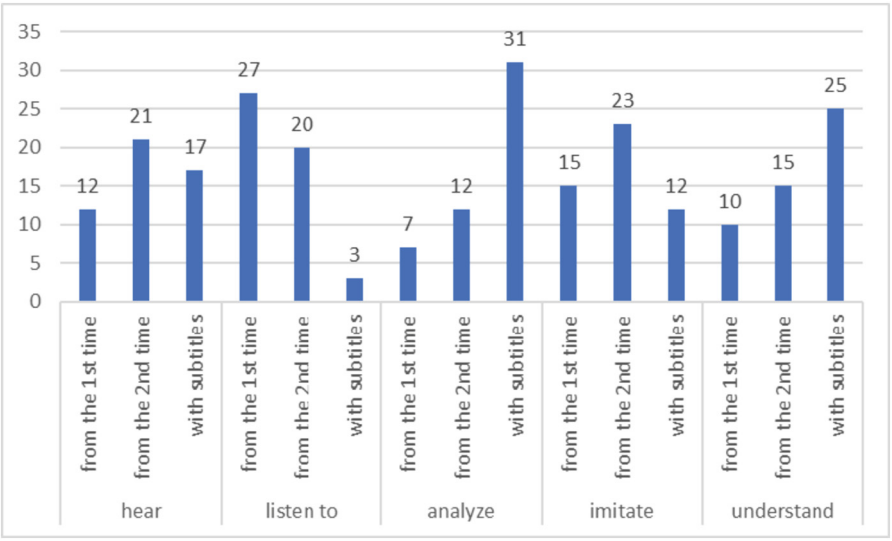


Fig. 2. The use of the one of distance learning methods on the example of video material «Transistors, how do they work».

This principle of working with video can be used in mixed learning. For example, it can be used during off-line English classes with a group of students to better master the material on the topic.

As alternative ways to work with video materials when learning a foreign language, we can offer the following methods.

Video materials can be used as a stage of self-preparation for a lesson on a specific topic. Students are given the task to watch and independently analyze the video at home. In addition to this video, students receive exercises to the video and a list of key vocabulary, which must also be completed and learnt at home. Thus, a teacher sets

the students on a certain topic and gives them the opportunity to learn the necessary vocabulary in advance. An off-line lesson can be started immediately with a discussion of the video material viewed home. The discussion should take place using the learned vocabulary. And in the course of the lesson, the teacher can gradually complicate the material with the help of new texts, videos and exercises.

Here is the next method. It combines work with text and video. Reading and translation of each text fragment alternates with watching a part of the video clip devoted to the same aspect as the previously read text fragment. Working with a text fragment should not exceed seven to ten minutes, and the duration of the video fragment should not exceed three to five minutes. After each such set (reading a text and watching a video), a teacher asks questions about the topic read and viewed. Thus, different types of activities alternate: monotonous reading and translation of a text, with watching a dynamic video; besides, information is fragmented. As a result of such method the studied material is activated in the process of reading, listening and speaking [6].

3 Results

The conducted experimental work proves that the use of mixed forms of learning, namely the introduction of on-line methods of teaching and learning a foreign language with the use of information and communication technologies has a positive effect on the results of students' progress [7, 8]. While conducting this study, it was practically proved that students in groups where learning takes place using both on-line and off-line methods, master new material and new vocabulary faster and more efficiently than students whose learning takes place using only traditional methods. Video-assisted learning has a positive effect on students of technical specialties, since the juxtaposition of video and text gives a complete understanding of the educational material.

4 Conclusion

It can be concluded that distance learning technologies are one of the most convenient, simple and affordable way to learn English. Thanks to such technologies, it is possible to significantly reduce training costs and achieve good results in teaching English. However, when receiving basic higher education, a combination of distance and traditional forms of education may be an ideal option for studying. Student performance depends on many factors, including the quality of educational programs, teaching methods, the level of students' motivation, individual needs and abilities of students, and many other factors. We hope that our study will bring practical benefits to teachers and students in building effective education for modern youth.

References

1. Kitaygorodskaya, G.A.: Intensive training in foreign languages. Theory and practice. Vysshaya shkola, the School of Kitaygorodskaya, Moscow, Russia (2014)

2. Kamardina, Y.S.: Web-quest technology in English lessons. In: Proceedings of the Thirteenth International part-time scientific and methodological conference "Innovative strategies for the development of teacher education", vol. 1, pp. 166–168. Center "Prosveshchenie", Saratov, Russia (2017)
3. Kamardina, Y.S.: Technology for developing a "web-quest" for teaching English. Discurs **6**(32), 90–96 (2019)
4. Kuznetsova, M.A.: Studying English in the conditions of distance learning. Int. J. Human. Natl. Sci. **12–2**(51), 191–194 (2020). https://doi.org/10.24411/2500-1000-2020-11570
5. Monakchov, V.M.: Informatization of educational and methodological support of the integral process of formation of competencies and technological monitoring of their quality management. Vestnik MSU **4**, 46–59 (2012)
6. Kappusheva, I.Sh., Erkenova, D.I., Nanay, F.A., Filipskaya, A.V.: Techniques for professional vocabulary activation for students with mosaic thinking. In: Proceedings of the International scientific and practical conference "Language discourse in social practice", pp. 195–200. TSU, Tver, Russia (2023)
7. Shchukin, A.N.: Modern intensive methods and technologies for teaching foreign languages. Filomatis, Moscow, Russia (2014)
8. Pikhart, M., Al-Obaydi, L.H.: Potential pitfalls of online foreign language teaching from the perspective of the university teachers. Heliyon **9**, 45–51 (2023)

Evaluating the Effectiveness of Flipped Classrooms Using Linear Regression

Roman Tsarev[1,2(✉)], Biswaranjan Senapati[3], Shadia Hamoud Alshahrani[4], Alsu Mirzagitova[5], Shokhida Irgasheva[6], and Joel Ascencio[7]

[1] MIREA - Russian Technological University (RTU MIREA), Moscow, Russia
tsarev.sfu@mail.ru
[2] Bauman Moscow State Technical University, Moscow, Russia
[3] University of Arkansas at Little Rock, Arkansas, IL, USA
[4] Medical Surgical Nursing Department, King Khalid University, Khamis Mushate, Saudi Arabia
[5] Kazan Federal University, 89, Kazanskaya Street, Yelabuga, Republic of Tatarstan 423604, Russia
[6] Tashkent Institute of Finance, Tashkent, Uzbekistan
[7] Universidad César Vallejo, Lima, Perú

Abstract. E-learning expands the potential of traditional education both in terms of technical, software and information tools, as well as in terms of the use of different ways, methods and approaches in learning. This paper examines student learning based on the flipped classroom model. The use of flipped classrooms involves students independently studying or repeating theoretical material before practical classes. This allows to focus on the practical aspects of the discipline being taught, as well as to develop skills and abilities based on the knowledge gained during lectures and independent flipped classroom preparation. The use of an electronic learning environment that contains the educational material needed to study the course makes the flipped classroom application technically feasible. This paper presents the results of applying linear regression to determine the relationship of student achievement to the time of learning theoretical material in an e-learning course. We used the least squares method for linear regression. The results of the research showed that flipped classrooms can free up time in a practical class from repetition or study of theoretical material, improve the efficiency of the learning process and student performance. #COMESYSO1120.

Keywords: E-Learning · Flipped Classroom · Linear Regression · Least Squares Method · LMS · Moodle

1 Introduction

Electronic learning opens up new opportunities in education [1–4]. When organizing e-learning, innovative educational technologies are used, which automate and optimize the learning process, implement the connection between theory and practice without reference to the location, presence in the classroom, increase the efficiency of time management of students and teacher [5–9].

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 418–427, 2024.
https://doi.org/10.1007/978-3-031-53552-9_38

Modern information and communication technology provides a variety of means for effective organization and support of learning, improving the level of students' knowledge [10]. A special place in e-learning is taken by students' independent work, since classroom time is often not enough, in particular for the application of the acquired knowledge in practice.

Independence is defined by teacher-researchers as the ability of an individual to carry out educational activities without interference from outside, to be guided by their personal experience. Independence is expressed in the ability to set goals for themselves and achieve them by their own efforts. It is the development of this skill at a high level will help the student to make decisions under conditions of uncertainty and risk in the process of practice.

The issue of organizing students' independent work remains one of the most urgent nowadays [11]. Teachers are faced with the inability and unwillingness of students to effectively organize their extracurricular study time, it especially concerns students who fail. Therefore, this form of work requires a high level of motivation of students, consciousness, and active participation in learning [12].

In order to encourage students to work independently, in our opinion, it is worthwhile to devote extracurricular time to studying theory, so that students would be interested in obtaining or updating theoretical knowledge, without which they would not be able to perform practical tasks, because there would be no more time allocated for this in class as part of practice.

Successful study of theoretical and practical material by students plays a crucial role in applying the acquired knowledge and skills in solving practical knowledge-intensive tasks in the future [13–15]. Therefore, it is essential for students to form key competencies that will enable them to solve professional tasks after graduation.

It is not uncommon for the workload in the curriculum to be distributed in such a way that more time is devoted to theory, while half as much time may be devoted to practice.

The practice of teaching shows that by the time of practical classes some of the theory is forgotten by students, it is a natural process, but it makes it impossible to solve practical problems immediately [16]. Thus, the teacher has to repeat theoretical material first, which takes time away from practical assignments. In this regard, it is recommended to build the learning process on the principle of flipped classrooms.

The method of flipped classrooms is a form of blended learning, different from the traditional approach, in which the classroom and extracurricular work swap places [17–20]. In this approach, the process of knowledge transfer is fully or partially shifted to independent learning [21–23].

The main task of a teacher who educates according to the flipped classroom principle is to increase students' motivation to search for knowledge independently outside the classroom, develop such qualities as self-control and self-organization, because the time spent on independent studying and working on theoretical material directly depends on the success in practical assignments [24–27].

Some principles of this method were used about twenty years ago by American teachers who provided new theoretical material in the format of video classes for self-study and it helped to decrease significantly the number of failing students. Teachers

suggest paying attention to what kind of lesson should be "flipped," then choose the type of independent work, the means of learning, electronic resources, and pay attention to group forms of work.

Currently, with the active use of e-learning tools, the flipped classroom method has gained particular popularity among teachers of higher education institutions.

Researchers distinguish three models of flipped classes: classical, advanced, and systemic or combined model. The classical model assumes that the student is familiarized in advance with the theoretical material of the upcoming class (independent study of theory and practical application of knowledge in class). Materials for independent studying may be lectures, presentations, literature offered by the teacher, video and audio documents.

The advanced model of flipped learning includes two stages of extracurricular (independent implementation of information search) and classroom (discussion and analysis of search results) work. Students can work not only individually but also in pairs or groups. As part of extracurricular work, students prepare theses, discussion questions, then place the results in the used electronic educational environment, so that the teacher can monitor the independent work of each learner and other students in advance to familiarize with the material and prepare. At the classroom stage, students present their results, discuss and analyze the material studied, put forward hypotheses and arguments, organize debates or round tables.

The system model of the flipped classroom combines the content of the previous two models. A distinctive feature of this model is, first of all, the practical application of theory and only then its theoretical justification. That is, students first work independently with a problematic situation or task, offer options for its solution by searching and analyzing the necessary information, and then, under the guidance of the teacher, discuss and analyze the proposed solutions. However, the stages of work do not end there: the next step is a repeated independent study of the theoretical basis of the issue in a distance format. At the final stage, as part of this model of flipped learning, the teacher and students summarize the results, identify the correct solution to the problem on the basis of the studied theoretical material.

Thus, the method of flipped classrooms intersects with problem-based and person-centered learning, students use different approaches to search, analyze, systematize and actualize theoretical information, independently organize and evaluate their learning activities, predict the results, competence-based as well as system-activity approach is provided. All this is an undeniable advantage of this approach.

Certainly, there are also a number of difficulties that the teacher has to face. They can be related to technical equipment or lack of necessary professional qualities of the teacher himself/herself, because preparation of extracurricular theoretical material is not an easy additional large-scale work of a teacher, it is important to structure the material qualitatively and not to forget about using interactive teaching methods, the principle of multimedia, which helps to increase students' motivation for learning.

In addition, the personal qualities and characteristics of students can also cause difficulties in the process of implementing flipped classroom. Many students need direct live contact with the teacher. This improves memorization, the distant format of classes does not always provide an opportunity to ask a question and instantly get an answer.

Each of these aspects needs to be taken into account by an instructor who teaches the flipped classroom method.

In our study we used the classical model of flipped classrooms, which was implemented in the LMS Moodle to confirm the effectiveness of this approach for practical learning of the material.

2 Methods

Moodle is a learning management system that allows to organize efficiently interaction between a teacher and students [28–31]. It contains theoretical materials: lecture notes, presentations, educational materials for self-study, which can be used for flipped classroom.

It is important that the LMS Moodle allows recording the time spent by the student on studying the material. In our work we evaluate the dependence of the results of the student's progress on the time spent on studying the educational material.

Moodle provides a wide range of testing options [30–33]. Moodle testing system is flexible, it allows to create questions of different types and formats, makes it possible to analyze test results of an individual student, a group or several groups.

In this paper, the evaluation of the dependence of testing results on the time spent by a student in an electronic course is performed using linear regression [34–36].

The linear regression problem for the two-dimensional case is that there are two variables x and y, and we need to find the relationship between them. That is, express one variable through the other using the linear function $y = f(x) = a \cdot x + \mathrm{b}$, where a and b are constants.

The possibility of constructing this dependence is determined by Pearson correlation coefficient [37]. The Pearson correlation coefficient is in the interval $[-1; 1]$. If the value of Pearson correlation coefficient for the obtained sample belongs to the range $[-0.25; 0.25]$, it means that the correlation is very low, there is practically no dependence between the variables and it makes no sense to apply linear regression. If the Pearson correlation coefficient belongs to the interval $[-1; -0.75] \cup [0.75; 1]$, then the linear correlation is very high and the dependence can be found by applying, for example, the least squares method. If the value of the Pearson correlation coefficient belongs to the interval $[-0.75; -0.25] \cup [0.25; 0.75]$, then there is a relationship, but a weak one; linear regression can be applied at the researcher's discretion. Pearson correlation coefficient can be calculated using the following formula:

$$\frac{\sum_{i=1}^{n} x_i y_i - n\overline{x}\overline{y}}{\sqrt{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2}\sqrt{\sum_{i=1}^{n} y_i^2 - n\overline{y}^2}},$$

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i.$$

The linear regression problem can be solved by the basic method of regression analysis - by the method of least squares. Having data sets (x_i, y_i), $i \in [1, n]$, where n is

the number of points, let us calculate the dependence of y on x: $y = f(x)$, where function $f(x)$ is linear:

$$y = f(x) = ax + b,$$

$$a = \frac{\sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i \right)^2},$$

$$b = \frac{n \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i \right)^2}.$$

The problem can be stated as follows:

$$\sum_{i=1}^{n} (y_i - f(x_i))^2 \to \min.$$

That is, it is necessary to choose such coefficients a and b in the function $f(x)$ that the sum of the squares of the subtraction of the already known value of the variable y and the obtained value of the function from the corresponding variable x is minimal. In this paper, the obtained line shows the dependence of students' performance on the time spent studying the educational material.

3 Results and Discussion

For the experiment described, the study was conducted on five groups of students. The groups ranged in size from 28 to 32 students. The dependence of the test scores on the time spent in the electronic course is shown in Fig. 1.

Note that it was not the total time that was taken into account, but the time during which the student studied the theoretical material.

For this experiment the Pearson correlation coefficient was equal to 0.884. The dependence of the variables is described as follows:

$$\mathrm{y} = 0.198 \cdot \mathrm{x} + 2.962.$$

Thus, the results show that the more time a student spent in a course, the better his or her performance. This approach gives an appropriate representation on a medium to large sample of data. The approach shows a general trend.

Certainly, there are students who, after spending a lot of time in an e-course, nevertheless perform poorly on tests. In contrast, there are students who score high despite spending relatively little time in the e-course. In this case, it is the personal qualities of individual students that are at play.

However, on the whole, the proposed approach shows an objective degree of mastering the educational material, despite these particular deviations.

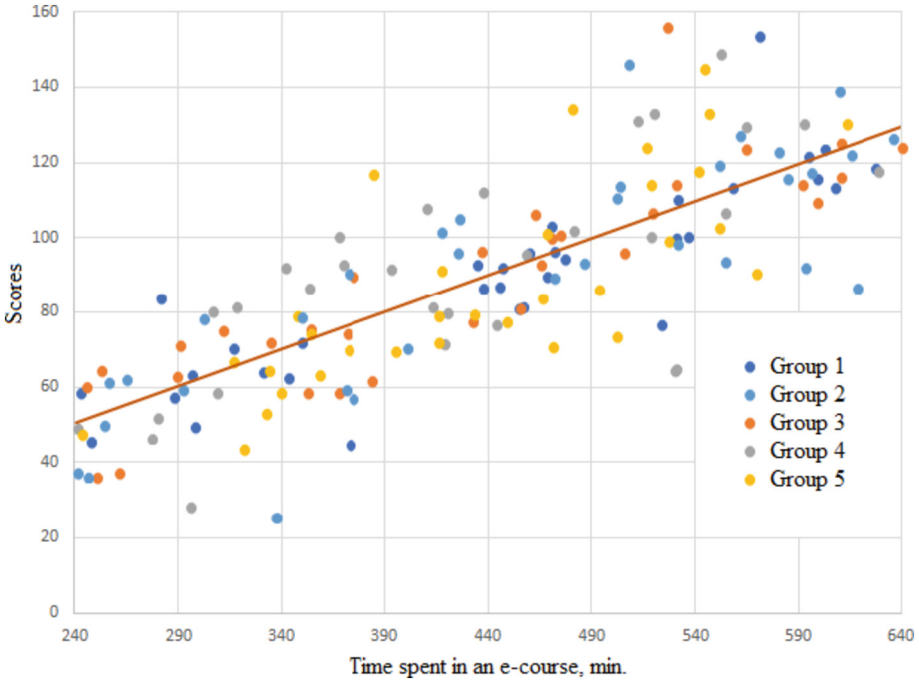


Fig. 1. Dependence of students' academic performance on the time of studying theoretical material in the e-course.

4 Conclusion

The modern education system is characterized by active use of various methods, forms and means of e-learning. The high labor intensity of the teacher's work, the shortage of teaching time devoted to practical assignments, encourages teachers to look for new, non-traditional ways of learning, which could optimize the learning process, motivate students to independence, self-organization and achievement of high results, to implement an effective time management.

The flipped classroom method makes it possible to reformat the curriculum and divide the learning material so that theory is allocated for students' independent work and practice is carried out directly in the classroom, this allows for problem-based, blended and individualized learning.

Flipped classrooms imply a change of roles: the teacher becomes a consultant, an assistant, an organizer, and students, in turn, become the center of the educational process, they independently determine the pace of learning, set certain goals for themselves.

Flipped classrooms enhance students' motivation. Assignments provided for independent work are always associated with scores: success in practical assignments depends on how much time a student spends studying theory and how good and effective this activity was.

The choice of flipped classroom is closely connected with the professional competence, skill, and progressiveness of the teacher. The teacher himself/herself should

be distinguished by creativity, be interested in achieving high results, master new pedagogical methods, create multimedia content. The way and means by which he/she will implement extracurricular work with theory, structure educational material is of great importance for all subjects of educational process. To the most effective methods of e-learning can be attributed video lessons, presentations, including quizzes, game presentations, various additional video materials, webinars, podcasts and forums.

So, the approach of flipped classroom provides greater involvement of students in the learning process, promotes formation of a dynamic, creative, cognitive environment, develops critical thinking and creativity. Students take more responsibility for their studying and learning results. Knowledge is not given in ready-made form, it requires search, comprehension, cognition on the part of the students.

Extracurricular work with theoretical material is focused on all students: those who need additional processing of the received knowledge at their own pace, learners who actualize theoretical material to prepare for the practical session, as well as those students who may have missed classes in the discipline and can study theory independently without presence in the classroom, at a comfortable pace and place, get an equal amount of knowledge and opportunities for preparation.

The flipped classrooms promote the activation of semantic, non-mechanical memory. This ensures the quality and duration of the learning outcome.

Summarizing about the obtained results of the study, we can note the fact that the spent and recorded time of students in the electronic course created in Moodle is reflected in the performance of students. It can be observed that the active time of the learner in the electronic environment is related to his/her preparation and study of theoretical material. The results of the linear regression method show that the more time the student spent in the course, the better his/her performance.

As a further study, the use of the system model of the flipped classroom, identifying the most effective methods of independent work can be proposed.

References

1. An, H., Chen, J.: ElearnChain: a privacy-preserving consortium blockchain system for e-learning educational records. J. Inform. Secur. Appl. **63**, 103013 (2021). https://doi.org/10.1016/j.jisa.2021.103013
2. Bhardwaj, P., Gupta, P.K., Panwar, H., Siddiqui, M.K., Morales-Menendez, R., Bhaik, A.: Application of Deep Learning on Student Engagement in e-learning environments. Comput. Electr. Eng. **93**, 107277 (2021). https://doi.org/10.1016/j.compeleceng.2021.107277
3. Lin, H.-M., et al.: A review of using multilevel modeling in e-learning research. Comput. Educ. **198**, 104762 (2023). https://doi.org/10.1016/j.compedu.2023.104762
4. Maulana, F.I., Febriantono, M.A., Raharja, D.R.B., Khaeruddin, Herasmara, R.: Twenty years of e-learning in health science: a bibliometric. Procedia Comput. Sci. **216**, 604–612 (2023). https://doi.org/10.1016/j.procs.2022.12.175
5. Al Ghawail, E.A., Yahia, S.B.: Using the E-learning gamification tool Kahoot! to learn chemistry principles in the classroom. Procedia Comput. Sci. **207**, 2667–2676 (2022). https://doi.org/10.1016/j.procs.2022.09.325
6. Ipanaqué-Zapata, M., Figueroa-Quiñones, J., Bazalar-Palacios, J., Arhuis-Inca, W., Quiñones-Negrete, M., Villarreal-Zegarra, D.: Research skills for university students' thesis in E-learning: scale development and validation in Peru. Heliyon **9**(3), e13770 (2023). https://doi.org/10.1016/j.heliyon.2023.e13770

7. Liu, Y.: Matches and mismatches between university teachers' and students' perceptions of E-learning: a qualitative study in China. Heliyon **9**(6), e17496 (2023). https://doi.org/10.1016/j.heliyon.2023.e17496
8. Neffati, O.S., et al.: An educational tool for enhanced mobile e-Learning for technical higher education using mobile devices for augmented reality. Microprocess. Microsyst. **83**, 104030 (2021). https://doi.org/10.1016/j.micpro.2021.104030.технологии
9. Tsarev, R., et al.: Gamification of the graph theory course. Finding the shortest path by a greedy algorithm. In: Silhavy, R., Silhavy, P. (eds.) Networks and systems in cybernetics. CSOC 2023. LNNS, vol. 723. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-35317-8_18
10. Deetjen-Ruiz, R., et al.: Applying ant colony optimisation when choosing an individual learning trajectory. In: Silhavy, R., Silhavy, P. (eds.) Networks and systems in cybernetics. CSOC 2023. LNNS, vol. 723. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-35317-8_53
11. Gushchin, A.N.: Design-based science: curriculum for architects. AIP Conf. Proc. **2657**, 020047 (2022). https://doi.org/10.1063/5.0107174
12. Gushchin, A.N., Divakova, M.N.: Nurturing cognitive skills in undergraduates with the help of ontological analysis. IOP Conf. Ser. Mater. Sci. Eng. **972**, 012059 (2020). https://doi.org/10.1088/1757-899X/972/1/012059
13. Senapati, B., Talburt, J.R., Bin Naeem, A., Batthula, V.J.R.: Transfer learning based models for food detection using ResNet-50. In: Proceedings of the 2023 IEEE international conference on electro information technology (eIT), pp. 224–229. IEEE, Romeoville, IL, USA (2023). https://doi.org/10.1109/eIT57321.2023.10187288
14. Senapati, B., Rawal, B.S.: Adopting a deep learning split-protocol based predictive maintenance management system for industrial manufacturing operations. In: Hsu, CH., Xu, M., Cao, H., Baghban, H., Shawkat Ali, A.B.M. (eds.) Big data intelligence and computing. DataCom 2022. LNCS, vol. 13864. Springer, Singapore (2023). https://doi.org/10.1007/978-981-99-2233-8_2
15. Soomro, A.M., et al.: In MANET: an improved hybrid routing approach for disaster management. In: Proceedings of the 2023 IEEE international conference on emerging trends in engineering, sciences and technology (ICES&T), pp. 1–6. IEEE, Bahawalpur, Pakistan (2023). https://doi.org/10.1109/ICEST56843.2023.10138831
16. Nácher, M.J., Badenes-Ribera, L., Torrijos, C., Ballesteros, M.A., Cebadera, E.: The effectiveness of the GoKoan e-learning platform in improving university students' academic performance. Stud. Educ. Eval. **70**, 101026 (2021). https://doi.org/10.1016/j.stueduc.2021.101026
17. Chen, T., Luo, H., Wang, P., Yin, X., Yang, J.: The role of pre-class and in-class behaviors in predicting learning performance and experience in flipped classrooms. Heliyon **9**(4), e15234 (2023). https://doi.org/10.1016/j.heliyon.2023.e15234
18. Flores-Alarcia, Ò., del-Arco Bravo, I., Ramos-Pla, A., Espart, A.: Without fear of change: the flipped classroom as a flexible model in different learning environments. Heliyon **8**(12), e12160 (2022). https://doi.org/10.1016/j.heliyon.2022.e12160
19. Ng, E.K.L.: Student engagement in flipped classroom in nursing education: an integrative review. Nurse Educ. Pract. **68**, 103585 (2023). https://doi.org/10.1016/j.nepr.2023.103585
20. Schmid, R.F., Borokhovski, E., Bernard, R.M., Pickup, D.I., Abrami, P.C.: A meta-analysis of online learning, blended learning, the flipped classroom and classroom instruction for pre-service and in-service teachers. Comput. Educ. Open **5**, 100142 (2023). https://doi.org/10.1016/j.caeo.2023.100142

21. Barrios, T.A., Rojas, S.L., Movilla, J.S., Ibáñez, S.F.U., Taboada, A.H.: Characterization of flipped classroom model in higher education: a perception from educational resilience during Covid-19 pandemic. Procedia Comput. Science **203**, 575–582 (2022). https://doi.org/10.1016/j.procs.2022.07.082
22. Lapitan, L.DS., Chan, A.L.A., Sabarillo, N.S., Sumalinog, D.A.G., Diaz, J.M.S.: Design, implementation, and evaluation of an online flipped classroom with collaborative learning model in an undergraduate chemical engineering course. Educ. Chem. Eng. **43**, 58–72 (2023). https://doi.org/10.1016/j.ece.2023.01.007
23. Wilson, K.E., Hobbs, J.R.: Innovative use of a flipped-classroom approach to teach fundamental nursing skills. Teach. Learn. Nursing **18**(1), 144–147 (2023). https://doi.org/10.1016/j.teln.2022.08.002
24. Nacaroğlu, O., Bektaş, O.: The effect of the flipped classroom model on gifted students' self-regulation skills and academic achievement. Think. Skills Creat. **47**, 101244 (2023). https://doi.org/10.1016/j.tsc.2023.101244
25. Nja, C.O., Orim, R.E., Neji, H.A., Ukwetang, J.O., Uwe, U.E., Ideba, M.A.: Students' attitude and academic achievement in a flipped classroom. Heliyon **8**(1), e08792 (2022). https://doi.org/10.1016/j.heliyon.2022.e08792
26. Ruiz-Jiménez, M.C., Martínez-Jiménez, R., Licerán-Gutiérrez, A., García-Martí, E.: Students' attitude: key to understanding the improvement of their academic results in a flipped classroom environment. Int. J. Manage. Educ. **20**(2), 100635 (2022). https://doi.org/10.1016/j.ijme.2022.100635
27. Wittmann, S., Wulf, T.: Effects of flipped classes on student learning: the role of positively perceived instructor attitude towards students. Int. J. Manage. Educ. **21**(1), 100735 (2023). https://doi.org/10.1016/j.ijme.2022.100735
28. Bengueddach, A., Boudia, C., Bouamrane, K.: Interpretive analysis of online teaching labs constructed using moodle during the pandemic period. Heliyon **9**(5), e16007 (2023). https://doi.org/10.1016/j.heliyon.2023.e16007
29. Hachicha, W., Ghorbel, L., Champagnat, R., Zayani, C.A., Amous, I.: Using process mining for learning resource recommendation: a moodle case study. Procedia Comput. Sci. **192**, 853–862 (2021). https://doi.org/10.1016/j.procs.2021.08.088
30. Romero, E., García, L., Ceamanos, J.: Moodle and Socrative quizzes as formative aids on theory teaching in a chemical engineering subject. Educ. Chem. Eng. **36**, 54–64 (2021). https://doi.org/10.1016/j.ece.2021.03.001
31. Tsarev, R., et al.: Improving test quality in e-learning systems. In: Silhavy, R., Silhavy, P. (eds.) Networks and systems in cybernetics. CSOC 2023. LNNS, vol. 723. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-35317-8_6
32. Aljarbouh, A., et al.: Application of the K-medians clustering algorithm for test analysis in E-learning. In: Silhavy, R., Silhavy, P., Prokopova, Z. (eds.) Software engineering application in systems design. CoMeSySo 2022. LNNS, vol. 596. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-21435-6_21
33. Hertsch, M.F.: Multimedia-based enrichment for foreign language teaching. Procedia. Soc. Behav. Sci. **70**, 615–621 (2013). https://doi.org/10.1016/j.sbspro.2013.01.100
34. Boot, T.: Joint inference based on Stein-type averaging estimators in the linear regression model. J. Econ. **235**(2), 1542–1563 (2023). https://doi.org/10.1016/j.jeconom.2023.01.006

35. Stephan, T., Trappitsch, R.: Reliable uncertainties: error correlation, rotated error bars, and linear regressions in three-isotope plots and beyond. Int. J. Mass Spectrom. **491**, 117053 (2023). https://doi.org/10.1016/j.ijms.2023.117053
36. Toft, H.B., Müller, K., Hendrikx, J., Jaedicke, C., Bühler, Y.: Can big data and random forests improve avalanche runout estimation compared to simple linear regression? Cold Reg. Sci. Technol. **211**, 103844 (2023). https://doi.org/10.1016/j.coldregions.2023.103844
37. Edelmann, D., Móri, T.F., Székely, G.J.: On relationships between the Pearson and the distance correlation coefficients. Statist. Probab. Lett. **169**, 108960 (2021). https://doi.org/10.1016/j.spl.2020.108960

Building a Digital Twin of a Social Media User Based on Implicit Profile Data

Ilya Shirokov[1], Zulfiya Kamaldinova[1], Irina Dubinina[1], and Anton Ivaschenko[2(✉)]

[1] Samara State Technical University, Molodogvardeyskaya 244, Samara, Russia
[2] Samara State Medical University, Chapaevskaya 89, Samara, Russia
anton.ivashenko@gmail.com

Abstract. This presents an algorithm for analyzing public open data gathered from social media. The problem of extracting information from Big Data collected from various social networks is considered. The article presents a new software solution to build a digital twin for a social media user based on implicit information, determined by substituted parameters. The algorithm is implemented in Python and uses social network API. The process of data collection, their pre-processing and filtering, as well as methods for analyzing the received data are described. The effectiveness of the algorithm is demonstrated on the example of data analysis from the popular social network VKontakte. The results obtained allow us to conclude that the proposed algorithm can be used to analyze and extract useful information from social media. The paper is oriented for engineers in the field of data analysis and social media. #COMESYSO1120.

Keywords: Social Media · Implicit Profile Data · Digital Twin · Data Analysis · Big Data · Social System

1 Introduction

Modern social media analysis allows one to study interactions between people and their network communication, predict their behavior, classify and model information flows, etc. Thanks to the development of computer technology, people have been able to communicate virtually through social platforms. Analysis of information and behavior of people in the process of communication in this format is faster and more convenient. That is why the analysis of computer social networks is of great interest to modern researchers and is an important tool for understanding current trends in society.

Analyzing this data provides valuable insight into social interactions, user interests and behavior. In this paper, we provide an algorithm for analyzing open data from the social network VKontakte. The focus is on obtaining information about the city and age of the user. We propose an algorithm that uses the VKontakte API to extract data about users and their friends, and then analyzes this information to determine the user's city and age. The proposed algorithm can be useful for researching user behavior in social networks and for developing tools that use data from social networks.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 428–433, 2024.
https://doi.org/10.1007/978-3-031-53552-9_39

2 State of the Art

In the modern world, people use social media to meet and communicate, exchange information with each other, and spend their leisure time. At the same time, users of the social network post information about themselves, their friends and relatives. By analyzing this data, companies can provide various services, advertising, and obtain statistics on various issues [1].

By analyzing user information obtained from social networks, you can evaluate its various features. Social media analysis provides understanding the problems facing today's social scientists [2, 3]. One of the challenging research directions in this area is concerned with study of social and psychological trends that influence economics [4].

In order to compile a full-fledged profile, it is necessary to study the psychological characteristics of the user [5], but often not all questionnaire fields are filled out. However, do not despair, as the missing data can be obtained from other social networks or through the user's social circle. It is required to restore the missing key elements of the user's meta-profile, such as full name, age and city of residence, they can significantly expand the understanding of her/him on the network and improve the quality of personalized interaction using information from the social network.

Thus, to conduct such an analysis, it is necessary to be able to identify and process indirect or implicit indicators of the activity of social network users [6, 7]. This information remains open data for others, but is not specified directly in the user profile. Processing this information requires development of specific algorithms.

Results of implicit social media data analysis can be represented in the form a model of a typical user of social media, by analogue with a "digital twin" [8, 9]. The concept of a digital twin is widely used in various problem domains like manufacturing or social systems study.

As part of this study, a review of several software solutions designed for the analysis of social networks was carried out. There are multiple software solutions available on the market. Brandwatch (https://www.brandwatch.com/) is a social media monitoring platform that allows you to track mentions of a brand and its competitors, analyze trends and determine audience sentiment.

Such solutions as Hootsuite _ Insights (https://www.hootsuite.com/), Talkwalker (https://www.talkwalker.com/), Mention (https://mention.com/en/), NetBase Quid (https://netbasequid.com/), and Socialbakers (https://emplifi.io/) are social media monitoring and analytics tools that allows you to track brand, competitor and industry mentions as well as analyze the performance of companies in social networks.

To check the accuracy of the algorithms in this area, the cross-validation method is used [10], which is an effective tool for assessing the generalization ability of algorithms. This method consists of splitting the original data into several non-overlapping subsets called folds.Then, at each step, one of the folds is used as a test set. Thus, we can evaluate the accuracy of the model on independent data samples [11].

Although the solutions mentioned above have a high degree of accuracy, they are not focused on the Russian social media market and do not provide analytics and decision-making support for social systems management. The proposed approach is perspective to be used in practice of social systems analysis and decision-making support [12].

3 Methods

To solve the problem of restoring implicit values of user meta-profile parameters from social networks, it is necessary to formalize the problem of data aggregation. To do this, you can use a matrix, where the columns represent the sources of information about the user, such as the user's friends, and the rows contain the values of meta-profile parameters, such as year of birth or hometown, obtained from social environment accounts.

Thus we have the following matrix.

$$\begin{array}{c} \\ a_1 \\ a_2 \\ \vdots \\ a_m \end{array} \begin{array}{c} \begin{array}{cccc} U_1 & U_2 & \cdots & U_n \end{array} \\ \begin{pmatrix} S_{11} & S_{12} & \cdots & S_{1n} \\ S_{21} & S_{22} & \cdots & S_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ S_{m1} & S_{m2} & \cdots & S_{mn} \end{pmatrix} \end{array}, \tag{1}$$

where $U_1 \ldots U_n$ - users who are in the social environment of the analyzed user. Attributes $a_1 \ldots a_m$ - components of the user's meta-profile, representing the main characteristics, such as age, gender, education, interests, etc. Each element of the matrix $S_{i,j}$ represents the value of a specific parameter of the user's meta-profile, obtained from his account in the social network.

This formalized model can be used to describe a typical or specific model of a social media user digital twin. However, not all users provide complete information about themselves, so if any parameter is missing, the corresponding matrix element is set to zero.

The principle of the algorithm is to analyze user data. In our case, we get a list of friends and education data. Accordingly, users with the specified educational institution were taken for the sample.

The central element of the calculations is a formula that allows you to get the exact values of the desired parameters.

$$CV\left(\mu, X^L\right) = \frac{1}{N} \sum\nolimits_{n=1}^{N} Q_n\left(\mu\left(X_n^m\right), X_n^k\right), \tag{2}$$

where mapping μ connects a selection of cases X^m with a function (algorithm) $a : X \rightarrow Y$. The quality functional $Q_n = Q_n\left(a_n, X_n^k\right)$ is used for each partition n based on the algorithm $a_n = \mu\left(X_n^m\right)$.

4 Results

The developed software code is implemented by a Python script designed to make requests to the VKontakte API and analyze the received data.

This algorithm does not include validation of the users' friends' data, but this is not always necessary. At the moment, the hypothesis obtained by our algorithm is used to build a distribution according to the described parameters. The hypothesis is that the

depth of verification should be double. Although the depth of analysis can be set to any level, test results show that depth two provides a more accurate estimate than depth one.

Although a depth of three may give a more accurate result, the time cost increases to the order of $O(n^3)$, where n is the number of friends the users affected in this analysis have. This can significantly slow down the process, and therefore is often used only in cases where maximum accuracy of the results is required.

For example, let us consider an algorithm for restoring the user's age. It requires a special approach, since many users of the social media do not indicate the year of their birth. The usual analysis of friends may not give the correct result due to the low probability. Our approach is to use information about schools or universities provided by the users. This allows you to determine the year of birth with great accuracy. Also, you can find users among friends who attended the same school or university with the same graduation years and with a non-empty year of birth.

For this algorithm, search with depth one is also suitable, since additional checks will not give a significant increase in recognition accuracy. The algorithm is like the following:

1. Connect to the Social media API using an access_token,which is issued after the user authentication process.
2. Select randomly 80% of users from a given list.
3. For each user, request for information about his friends, including the name of the city where they live.
4. Determine the most frequent city among each user's friends.
5. Compare the received city value with the real city value specified for this user in the real_city list.
6. If the received value matches the real value, increment the c_count counter.
7. Store the resulting city value in the all_city list for later analysis.
8. For each user request information about their age.
9. Calculate the average age of all users from the sample.
10. Repeat the process n times, where n is the given value.

Sourec code structure consists of connecting the necessary libraries (requests, datetime, collections) and declaring the necessary variables (client_id, redirect_uri, scope, avg_c, avg_a, access_token, n, users_list, real_city, real_age). Then comes the user authentication process, which goes through the browser and requires the user to allow access to his account. After the authentication process, the code starts executing the main algorithm.

5 Discussion

In this study, the data of 40 users were analyzed, for whom the current city of residence and age were known. To form an experimental sample consisting of 40 accounts, 80%, that is, 32 accounts, were randomly selected.

After analyzing of 30 such samples, the average accuracy of the algorithm was obtained. For the city of residence, this figure ranged from 80% to 96%, and for age - from 30% to 41%. These values were calculated as the ratio of successful matches to each sample of 30 accounts. Then we calculate the totals.

Average accuracy scores obtained from data analysis were found to be 92% for city of residence and 35% for age determination. However, it is worth considering that when evaluating the data, only exact matches of the city and the number of full years of users were taken into account. But to perform various tasks, such as contextual advertising by age, user groups are used.

Therefore, we compiled a graph of the distribution of the dependences of the difference between the age determined by the algorithm and the real age (known to us in advance) on the number of matches (see Fig. 1). It shows the number of matches with the real age, taking into account the difference between the real and the age obtained during the program.

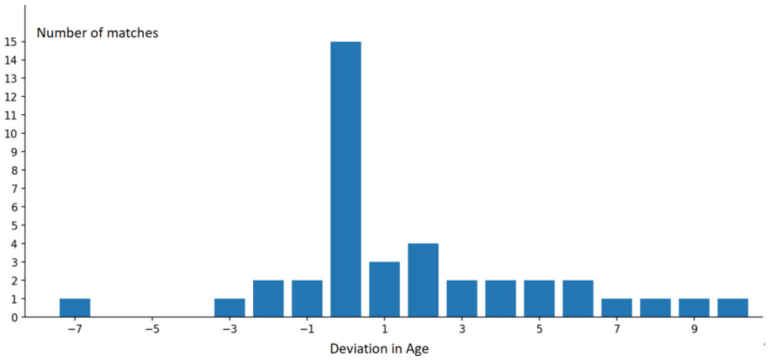


Fig. 1. Age deviation distribution.

It follows from it that the approximate age with an accuracy of three years, the age was correctly determined 29 times. Based on this, the age was calculated using cross-validation, taking into account the approximate age. The probability of correct identification increased to 68%. Therefore, by slightly increasing the age selection, it turned out to improve the accuracy of the algorithm.

6 Conclusion

The proposed approach allows building a digital twin of a social media user based on implicit profile data. Its application is illustrated by an example of VKontakte social media analysis to obtain information about the city and age of users. It was 92% for determining the current city of residence, 35% for determining the exact age and 72% for determining the approximate age with an error of three years.

Further research will include improving the accuracy of the estimate by adding various features to the program, e.g. analyzing data on other social media platforms in order to compare their characteristics and effectiveness in building digital twins, and studying other user parameters, such as gender, marital status, interests, etc., in order to expand the scope of the algorithms.

References

1. Zachlod, C., Samuel, O., Ochsner, A., Werthmüller, S.: Analytics of social media data – state of characteristics and application. J. Bus. Res. **144**(5), 1064–1076 (2022)
2. Chen, P.L., Cheng, Y.C., Chen, K.: Analysis of social media data: an introduction to the characteristics and chronological process. In: Big data in computational social science and humanities. Computational social sciences, pp. 297–321 (2018)
3. Batrinca, B., Treleaven, P.: Social media analytics: a survey of techniques, tools and platforms. AI & Soc. **30**, 89–116 (2014)
4. Darwiesh, A., Alghamdi, M., El-Baz, A., Elhoseny, M.: Social media big data analysis: towards enhancing competitiveness of firms in a post-pandemic world. J. Healthcare Eng. **6967158**, 1–14 (2022)
5. Brailovskaia, J., Teichert, T.: "I like it" and "I need it": Relationship between implicit associations, flow, and addictive social media use. Comput. Hum. Behav. **113**, 106509 (2020)
6. Yang, C., Tang, X., Dai, Q., Yang, H., Jiang, L.: Identifying implicit and explicit relationships through user activities in social media. Int. J. Electron. Commer. **18**, 73–96 (2013)
7. Villi, M., Matikainen, J.: Participation in social media: studying explicit and implicit forms of participation in communicative social networks. Media Commun. **4**, 109 (2016)
8. Ahlgren, J., et al.: Facebook's cyber–cyber and cyber–physical digital twins, pp. 1–9 (2021). https://doi.org/10.1145/3463274.3463275
9. Pigni, F., Watson, R., Piccoli, G.: Digital Twins: representing the future. SSRN Electron. J. (2021). https://doi.org/10.2139/ssrn.3855535
10. Kohavi, R.: A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Proceedings of the 14th international joint conference on artificial intelligence, pp. 1137–1145. Palais de Congres Montreal, Quebec, Canada (1995)
11. Langford, J.: Quantitatively tight sample complexity bounds. Carnegie Mellon Thesis, 124 p. (2002)
12. Golovnin, O., Ivaschenko, A., Krivosheev, A., Dodonova, E., Dubinina, I.: Socio-economic sustainability monitoring based on intelligent analysis of social media. Stud. Syst. Dec. Control **457**, 369–381 (2023)

Analysis of a Data Set to Determine the Dependence of Airline Passenger Satisfaction

V. S. Tynchenko[1,5,6], Borodulin[1], I. I. Kleshko[3(✉)], V. A. Nelyub[1,2], and Rukosueva[4]

[1] Artificial Intelligence Technology Scientific and Education Center, Bauman Moscow State Technical University, 105005 Moscow, Russia

[2] Peter the Great St. Petersburg Polytechnic University, Saint Petersburg, Russia

[3] Department of Information Economic System, Institute of Engineering and Economics, Reshetnev Siberian State University of Science and Technology, 660037 Krasnoyarsk, Russia
ilya-kleshko@mail.ru

[4] Department of Systems Analysis and Operations Research, Institute of Informatics and Telecommunications, Reshetnev Siberian State University of Science and Technology, 660037 Krasnoyarsk, Russia

[5] Information-Control Systems Department, Institute of Computer Science and Telecommunications, Reshetnev Siberian State University of Science and Technology, 660037 Krasnoyarsk, Russia

[6] Department of Technological Machines and Equipment of Oil and Gas Complex, School of Petroleum and Natural Gas Engineering, Siberian Federal University, 660041 Krasnoyarsk, Russia

Abstract. This paper presents an analysis of a data set to determine the factors influencing airline passenger satisfaction. The study examines various criteria such as gender, customer type, age, travel type, class, and range to assess their impact on passenger satisfaction. The dataset consists of 25 columns, including attributes like Wi-Fi availability, convenience of online booking, seat comfort, in-flight entertainment, baggage handling, and overall satisfaction. The sample is relatively balanced, with equal representation of men and women, predominantly repeat customers, and a majority flying for business purposes. Key findings include a strong correlation between departure and arrival delays, higher satisfaction among passengers in business class, and positive ratings for Wi-Fi service correlating with overall satisfaction. Correlation analysis reveals interdependencies between different attributes, such as the influence of cleanliness on seat comfort and food and beverage ratings. In addition, a neural network forecasting model is used to estimate the average ratings of passengers, although with low accuracy, which was later excluded. Finally, a decision tree algorithm is utilized to identify the most significant attributes affecting passenger satisfaction words.

Keywords: Data set analysis · airline passenger satisfaction · correlation analysis · neural network prediction · decision tree algorithm

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 434–458, 2024.
https://doi.org/10.1007/978-3-031-53552-9_40

1 Introduction

A data set represents processed and structured information in tabular form. The objects of such a table are rows, columns in turn represent attributes. Aggregate implies labeled data, which is based on machine learning. Big data is a specific term used to describe data sets so complex and large that they become difficult and inconvenient to work with.

This dataset looks at the ratings of airline passengers on different criteria based on factors such as gender, customer type, age of the passenger, type of travel, class, and range. The end result is an assessment of their satisfaction from the flight. Flying by airplanes is one of the most convenient and safe ways to get from one point to another, while pursuing conditions such as speed and good service. But despite the widespread development of this mode of travel, there are many things that may not seem modernized by the norms of the time. Then marketers use in turn special surveys to find out such things and then improve them [1–12].

2 Data Science

Data science focuses on extracting value from complex data sets. The complexity of different data lies in the variety of different sources from which it has been extracted. Data, given such a challenge, must in turn be organized in such a way that it can be properly interpreted and associated. Different approaches and methods establish causality, with proper investigation establishing meaningful connections between complex interactions in given systems. Thus, the more complex the data, the more complex will be the mathematical approaches including computation, machine learning and system-based methods [13–21].

3 Common Dataset

Surveys of consumers of certain things allow marketers to look at what they are willing to sell from different angles. This allows them to identify strengths and weaknesses in order to build on what they already have or correct deficiencies. There are a total of 25 columns in the database, including: id, gender, customer type, years on plane, travel type, class, distance, Wi-Fi, arrival/departure convenience, ease of online booking, cabin location, food and beverage, reservations, seat comfort, in-flight entertainment, in-flight service, leg room service, baggage handling, check-in service, onboard service, cleanliness, average score, departure delay in minutes, arrival delay in minutes, satisfaction [19, 22–30].

The sample is more or less balanced (56.7% versus 43.3%), which indicates the lack of a perfect balance and the need to generalize the results.

- The number of men and women in this sample is approximately equal;
- The vast majority of airline customers are repeat customers;
- Most customers were flying for business rather than personal reasons.
- About half of the passengers were in business class;
- More than 60% of passengers were satisfied with the baggage transportation service (rating 4–5 out of 5);

- More than 50% of passengers felt comfortable in their seats (score 4–5 out of 5);
- There was a strong correlation (96%) between the attributes "Departure delay in minutes" and "Arrival delay in minutes";
- The age range for most of the airline's regular customers is between 30 and 50 years old (with an average of just over 40 years old). The age range for non-permanent customers is slightly smaller (25 to 40 years old, with an average of just under 30);
- Customers whose range is long prefer to fly business class;
- The greater the distance traveled by an airplane passenger (respectively, the longer the flight), the more satisfied they are with in-flight entertainment and extra legroom (on average);
- The majority of passengers who flew economy plus or economy class were dissatisfied with the flight, while those who were lucky enough to fly business class were satisfied;
- Almost all passengers who rated the Wi-Fi service as 5 out of 5 were satisfied with the flight;
- The majority of passengers who rated seat comfort and extra legroom at 4 and 5 out of 5 were satisfied with the flight.

The raw data is presented in Fig. 1.

12 COL.1	12 id	ab Gender	ab Customer Type	12 years	ab Type of travel	ab Class	12 Flight distance	12 Wi...	1 Convenient departure/arrival time	
0	19556	Woman	Loyal Customer	52	Business	Eco	160	5		4
1	90035	Woman	Loyal Customer	36	Business	Business Class	2863	1		1
2	12360	Man	disloyal Customer	20	Business	Eco	192	2		0
3	77959	Man	Loyal Customer	44	Business	Business Class	3377	0		
4	36875	Woman	Loyal Customer	49	Business	Eco	1182	2		3
5	39177	Man	Loyal Customer	16	Business	Eco	311	3		3
6	79433	Woman	Loyal Customer	77	Business	Business Class	3987	5		5
7	97286	Woman	Loyal Customer	43	Business	Business Class	2556	2		2
8	27508	Man	Loyal Customer	47	Business	Eco	556	5		2
9	62482	Woman	Loyal Customer	46	Business	Business Class	1744	2		2
10	47583	Woman	Loyal Customer	47	Business	Eco	1235	4		1
11	115550	Woman	Loyal Customer	33	Business	Business Class	325	2		5
12	119987	Woman	Loyal Customer	46	Business	Business Class	1009	5		5
13	42141	Woman	Loyal Customer	60	Business	Business Class	451	1		1
14	2274	Woman	Loyal Customer	52	Business	Business Class	925	2		2
15	22470	Man	Loyal Customer	50	Personal	Eco	83	3		4
16	124915	Woman	Loyal Customer	31	Business	Eco	728	2		5
17	17836	Man	Loyal Customer	52	Personal	Eco +	1075	5		4
18	76872	Woman	Loyal Customer	43	Personal	Eco	1927	3		4
19	64287	Woman	Loyal Customer	50	Business	Business Class	3799	5		5
20	63995	Man	Loyal Customer	60	Business	Business Class	612	4		4
21	75855	Man	Loyal Customer	43	Personal	Eco	1437	3		4
22	106181	Man	Loyal Customer	55	Personal	Eco	302	1		2
23	44304	Man	Loyal Customer	25	Business	Business Class	1428	4		4
24	82602	Woman	disloyal Customer	30	Business	Eco	528	4		3
25	7823	Man	Loyal Customer	62	Personal	Eco	710	3		5

Fig. 1. Initial data

Correlation analysis was performed and the results are presented in Fig. 2. Correlation shows the interdependence of two or more variables chosen randomly. The meaning is that when the value of one variable changes, there is a regular change in the value of another variable [31–39].

The figure shows the most obvious dependencies, among which there are many obvious points: the higher the rating of baggage handling, the higher the service on board; the cleaner it was on board, the higher was the comfort of seats; also, the rating of food and drinks depended on cleanliness; as well as entertainment on board in turn also depended on food and drinks. Next, the neural network prediction in Fig. 3 was made. Correlation in this work is only a demonstration of dependencies and does not imply a causal relationship.

Satisfaction	Satisfaction	1
Ease of online booking	Wi-Fi	0,7106836255
Wi-Fi	Ease of online booking	0,7106836255
Purity	Entertainment on board	0,695268885
Entertainment on board	Purity	0,695268885
Purity	Seat comfort	0,6839244639
Seat comfort	Purity	0,6839244639
Purity	Food and drinks	0,6592533121
Food and drinks	Purity	0,6592533121
On-board service	Baggage handling	0,6319636119
Baggage handling	On-board service	0,6319636119
Entertainment on board	Food and drinks	0,6272645239
Food and drinks	Entertainment on board	0,6272645239
Entertainment on board	Seat comfort	0,6168172495
Seat comfort	Entertainment on board	0,6168172495
Seat comfort	Food and drinks	0,5809701314
Food and drinks	Seat comfort	0,5809701314
On-board service	On-board service	0,5547343663
On-board service	On-board service	0,5547343663
Class	Type of travel	0,5461841581
Type of travel	Class	0,5461841581
Baggage handling	On-board service	0,5249805705
On-board service	Baggage handling	0,5249805705

Fig. 2. Correlation analysis

	#	Average rating Forecast	Average rating	average error
2	1	2,966289427	1,214285714	0,005554
3	2	2,956250101	1,214285714	
4	3	2,955867063	1,214285714	
5	4	2,966194557	1,214285714	
6	5	3,605798498	1,285714286	
7	6	3,678626207	1,285714286	
8	7	3,641156891	1,285714286	
9	8	2,956910052	1,357142857	
10	9	2,964039214	1,357142857	
11	10	2,955963274	1,357142857	
12	11	2,966246376	1,357142857	
13	12	2,956520409	1,357142857	
14	13	3,60178086	1,357142857	
15	14	2,956751293	1,357142857	

Fig. 3. Forecast

This prediction gave an idea of the average ratings of passengers with low accuracy. The training of this network on the basis of the given estimates was excluded; later other methods will be used. To begin with, the decision tree in the figure was used to determine the most significant attributes in Fig. 4.

One of the most popular classification algorithms is the decision tree. It is a type of diagram that shows the possible consequences of decisions. Decision trees and rule-based expert systems are standard diagnostic tool.

Target Attribute: Satisfaction

№	Number	Attribute	Significance, %
1	13	Online booking	26,511
2	8	Wi-Fi	25,403
3	5	Type of travel	16,234
4	3	Customer Type	4,994
5	20	On-board service	4,288
6	21	Purity	3,563
7	19	Registration Service	3,073
8	14	Seat comfort	2,282
9	15	Entertainment on board	2,243
10	22	Average rating	1,805
11	1	id	1,714
12	11	Location of doors	1,676
13	6	Class	1,157
14	4	Years	1,149
15	16	On-board service	1,087
16	18	Baggage handling	1,010
17	7	Flight distance	0,783
18	17	Foot Room maintenance	0,534
19	10	Ease of online booking	0,275
20	9	Convenient departure time/approx...	0,113
21	12	Food and drinks	0,087
22	2	Paul	0,019

Fig. 4. Significance of attributes

The contiguity table showed a mean error of 3.04% of the total set, which is good in Fig. 5.

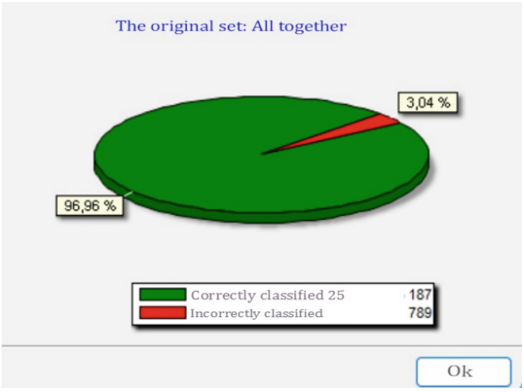


Fig. 5. Quality of classification

The output data was the level of satisfaction based on passenger ratings. Next, the neural network was trained, where the output data was the level of passenger satisfaction in Fig. 6.

Satisfaction	Content_out
Satisfied	Satisfied
Satisfied	Satisfied
neutral or unsatisfied	neutral or unsatisfied
Satisfied	Satisfied
Satisfied	neutral or unsatisfied
Satisfied	neutral or unsatisfied
Satisfied	Satisfied
Satisfied	Satisfied
Satisfied	Satisfied
Satisfied	Satisfied
Satisfied	neutral or unsatisfied
neutral or unsatisfied	neutral or unsatisfied

Fig. 6. Neural network result

In this case, the contiguity table shows an average error of 7.11%, Fig. 7.

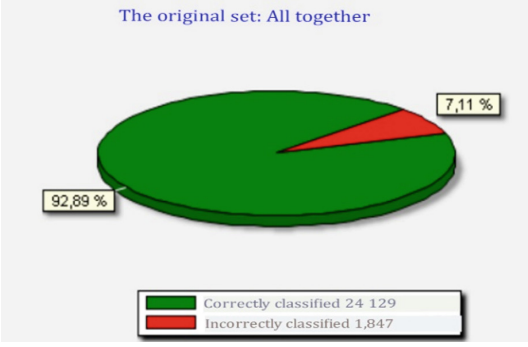


Fig. 7. Quality of classification

After that, another neural network was trained, but the output was the average rating (Fig. 8).

In this case, the neural network fails to draw adequate conclusions from the results. As the scatter plot in Fig. 9 shows, the neural network trained poorly.

Average Value_out	Average VALUE_ERR
4,14083536157429	2,9629054209524E-7
3,43720675695314	5,40514135007146E-6
2,17342673042716	6,77374024856007E-5
1,93408524382782	0,00159568508105863
2,57866800979109	0,00029865679912101
3,27901861343703	3,24965905563475E-6
4,50097375308605	0,00145919218929108
3,5176998150699	2,27084160000353E-5
3,65396309977192	8,94049040361188E-6
3,43667476138938	4,75966833371151E-6
2,7274368079597	0,000246178931695442
2,98044939871043	2,77057315402344E-5
4,56749137998743	0,00345183520387937
3,8695402478741	1,11406356817798E-5
3,59017822848585	2,54821495996595E-5
2,6009490285639	0,00012730504835912
3,15457133626775	9,94658856392683E-6
4,24802746131232	0,000102950504393751
3,86848996415555	9,33296626737959E-6
4,46465950887021	0,000826304731598854
4,36141867720004	0,000326871466798887
3,13916554379528	9,87821493298534E-7

Fig. 8. Neural network result

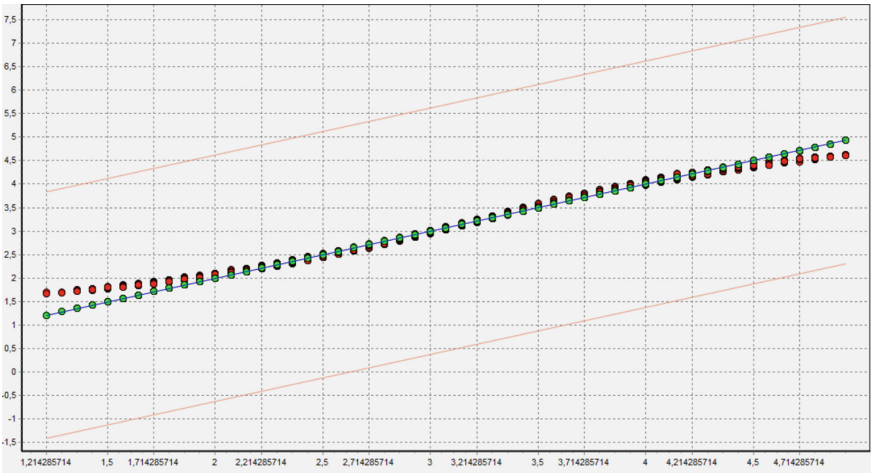


Fig. 9. Scatter diagram

4 Partitioning the Data Set into Groups

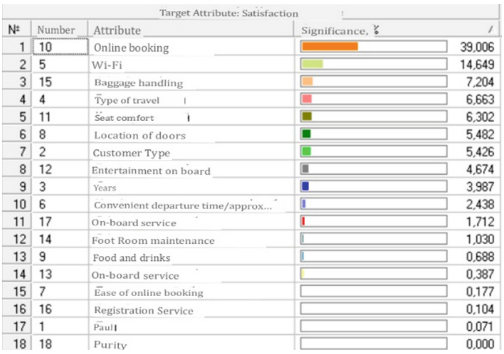
It was further identified that the data set needed to be categorized into meaningful groups in order to identify more dependencies.

This resulted in six groups, which were categorized by class, gender and type of travel.

This will allow you to better explore different groups with their own good opportunities to increase customer relevancy, although this may lead to the loss of certain dependencies in categories.

4.1 Class

In this group, the data were categorized into subgroups: business and eco. The business class subgroup was considered first. At the beginning, a decision tree construction procedure was carried out.



Target Attribute: Satisfaction

№	Number	Attribute	Significance, %
1	10	Online booking	39,006
2	5	Wi-Fi	14,649
3	15	Baggage handling	7,204
4	4	Type of travel	6,663
5	11	Seat comfort	6,302
6	8	Location of doors	5,482
7	2	Customer Type	5,426
8	12	Entertainment on board	4,674
9	3	Years	3,987
10	6	Convenient departure time/approx...	2,438
11	17	On-board service	1,712
12	14	Foot Room maintenance	1,030
13	9	Food and drinks	0,688
14	13	On-board service	0,387
15	7	Ease of online booking	0,177
16	16	Registration Service	0,104
17	1	Paul	0,071
18	18	Purity	0,000

Fig. 10. Significance of attributes

Figure 10 shows a table showing the significance of the attributes.

After that a neural network was built, with output data in the form of passenger satisfaction, Fig. 11 shows the contiguity table, which shows the good learnability of the neural network based on this data [41–49].

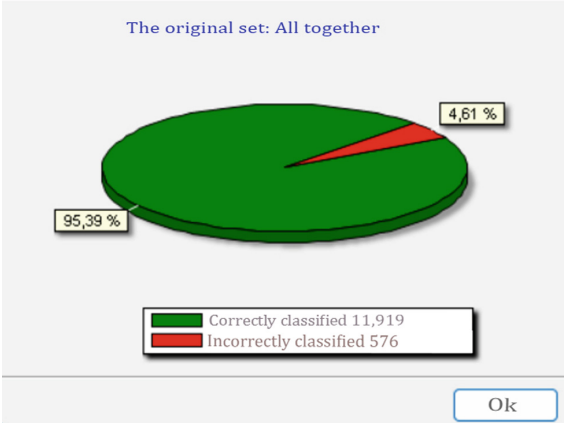


Fig. 11. Contiguity table

Clustering was then performed and the results are shown in Figs. 12, 13.

Fig. 12. Cluster profiles

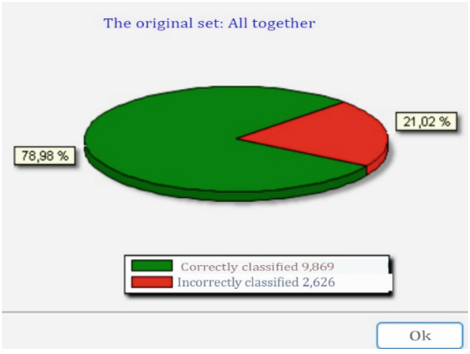


Fig. 13. Contiguity table

Next, the Kohonen map shown in Fig. 14 was constructed.

Graphical optimization is a standard method that is used to introduce the concept of optimization to practitioners.

SOM is a type of artificial neural network that uses a teacherless learning algorithm and has the additional property of preserving the topological mapping from the input space to the output space. Such that connected data points are mapped to neighboring positions in the map, preserving the relative distances between input data points. As a

result, it is an excellent tool for visualizing multidimensional data in lower dimensions [2, 12, 13].

Thus Kohonen maps are graphical optimization that helps to solve one- or two-dimensional optimization problems by drawing contours of the target and constraint functions. This visualization-based approach allows to understand and interpret the design variables and the behavior of the objective in terms of feasibility and optimality, allowing intuitive decision making. Visualization of higher dimensions presents higher complexity but is equally desirable. The Interpretable Self-Organizing Map (iSOM) is an artificial neural network that allows visualization of many dimensions using two-dimensional representations.

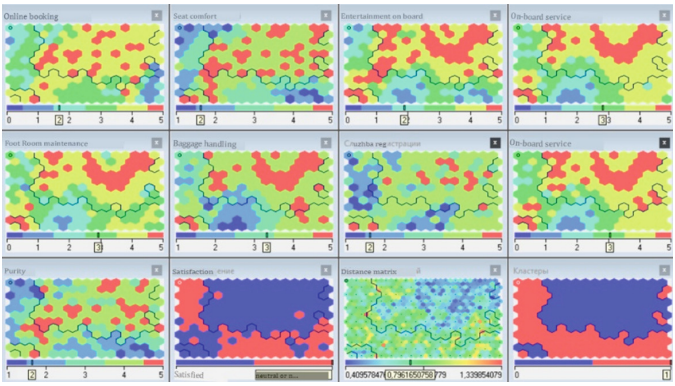


Fig. 14. Kohonen Map

After that, a subset of the economy class was considered.

At the beginning, the procedure of decision tree construction was carried out.

Target Attribute: Satisfaction

№	Number	Attribute	Significance, %
1	6	Wi-Fi	80,685
2	4	Type of travel	6,945
3	2	Customer Type	4,224
4	3	Years	1,742
5	18	On-board service	1,679
6	11	Online booking	1,447
7	16	Baggage handling	1,216
8	17	Registration Service	0,827
9	12	Seat comfort	0,523
10	14	On-board service	0,308
11	13	Entertainment on board	0,152
12	19	Purity	0,148
13	1	Paul	0,104
14	15	Foot Room maintenance	0,000
15	7	Convenient departure time/approx...	0,000
16	5	Class	0,000
17	8	Ease of online booking	0,000
18	10	Food and drinks	0,000
19	9	Location of doors	0,000

Fig. 15. Significance of attributes

Figure 15 shows a table showing the significance of the attributes.

After that a neural network was built, with output data in the form of passenger satisfaction, Fig. 16 shows the contiguity table, which shows the good learnability of the neural network on the basis of these data.

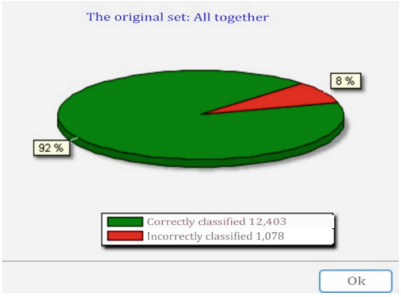


Fig. 16. Contiguity table

Clustering was then performed and the results are shown in Figs. 17, 18.

Fig. 17. Cluster profiles

Clustering is the task of grouping objects into clusters, which are in turn subsets, so that objects from one cluster are more similar to each other than objects from other clusters [17–20].

Clustering is an important data mining technique that is used in many real-world applications. A commonly used method to determine the number of clusters is to execute the clustering method several times, each time with a different number of clusters, and then select one of these clustering results according to a predefined target function, this is commonly referred to as the clustering confidence index [4].

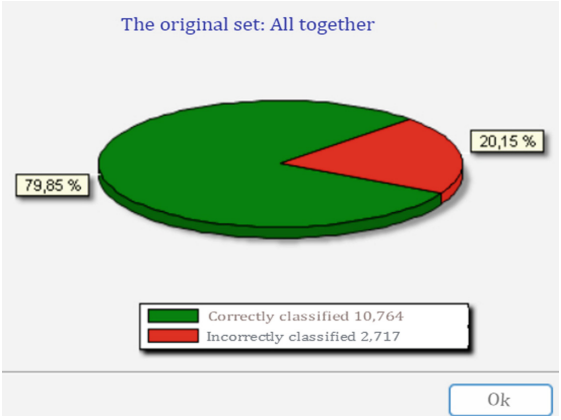


Fig. 18. Contiguity table

Next, the Kohonen map shown in Fig. 19 was constructed.

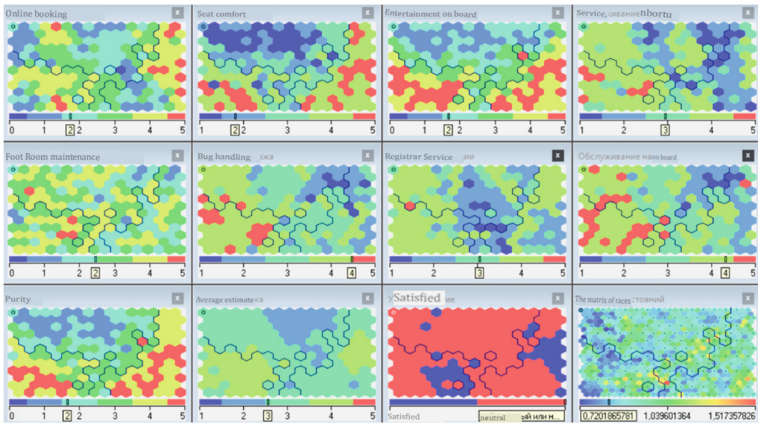


Fig. 19. Kohonen Map

In total, we can conclude that in the economy class the importance of such attribute as Wi-Fi was the highest than the other attributes. Neural network showed almost identical result. The power of clusters was weaker in the economy class subgroup. Kohonen maps showed an identical result.

4.2 Gender

A subgroup of women was considered first.

At the beginning, the procedure of building a decision tree was carried out.

Target Attribute: Satisfaction

N	Number	Attribute	Significance, %
1	10	Online booking	31,106
2	5	Wi-Fi	28,542
3	3	Type of travel	20,059
4	13	On-board service	3,734
5	4	Class	3,422
6	2	Years	2,700
7	1	Customer Type	2,650
8	15	Baggage handling	1,843
9	12	Entertainment on board	1,446
10	11	Seat comfort	1,177
11	6	Convenient departure time/approx...	1,143
12	8	Location of doors	1,053
13	9	Food and drinks	0,626
14	14	Foot Room maintenance	0,395
15	17	On-board service	0,059
16	16	Registration Service	0,044
17	7	Ease of online booking	0,000

Fig. 20. Significance of attributes

Figure 20 shows a table showing the significance of the attributes.

After that the neural network was built, with output data in the form of passenger satisfaction, Fig. 21 shows the contiguity table, which shows the good learning ability of the neural network based on this data.

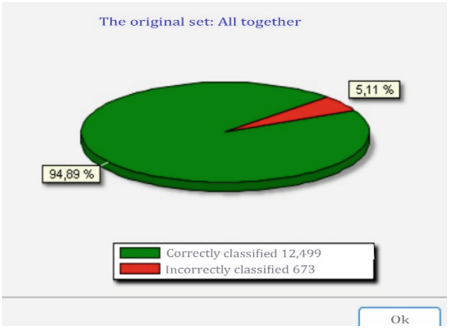


Fig. 21. Contiguity table

Clustering was then performed and the results are shown in Figs. 22, 23.

Fig. 22. Cluster Profiles

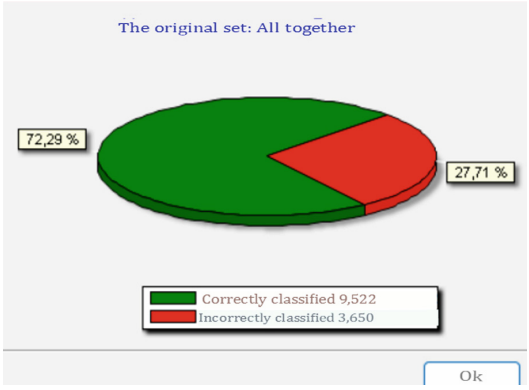


Fig. 23. Contiguity table

Next, the Kohonen map shown in Fig. 24 was constr.

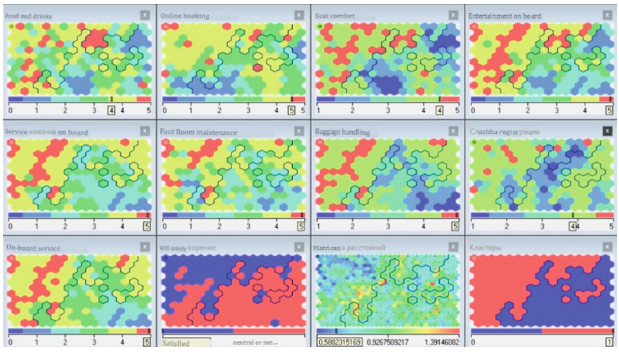


Fig. 24. Kohonen Map

A subset of men was then considered.
At the beginning, a decision tree construction procedure was carried out.

Target Attribute: Satisfaction

№	Number	Attribute	Significance, %
1	10	Online booking	34,166
2	5	Wi-Fi	22,820
3	3	Type of travel	21,642
4	1	Customer Type	3,865
5	15	Baggage handling	2,851
6	11	Seat comfort	2,806
7	2	Years	2,353
8	12	Entertainment on board	2,167
9	18	Purity	1,943
10	8	Location of doors	1,860
11	4	Class	1,099
12	14	Foot Room maintenance	0,614
13	16	Registration Service	0,613
14	17	On-board service	0,453
15	6	Convenient departure time/approx...	0,426
16	7	Ease of online booking	0,196
17	13	On-board service	0,126
18	9	Food and drinks	0,000

Fig. 25. Significance of attributes

Figure 25 shows the table showing the significance of attributes. After that a neural network was built, with output data in the form of passenger satisfaction, Fig. 26 shows the contiguity table, which shows the good learnability of the neural network based on this data.

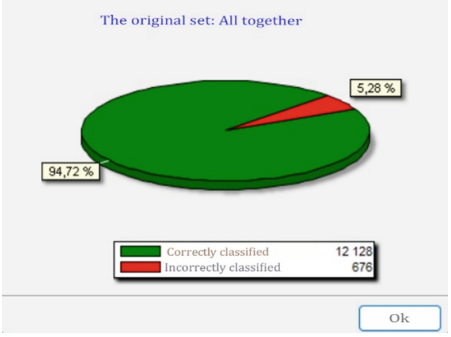


Fig. 26. Contiguity table

Clustering was then performed and the results are shown in Figs. 27, 28.

Fig. 27. Cluster profiles

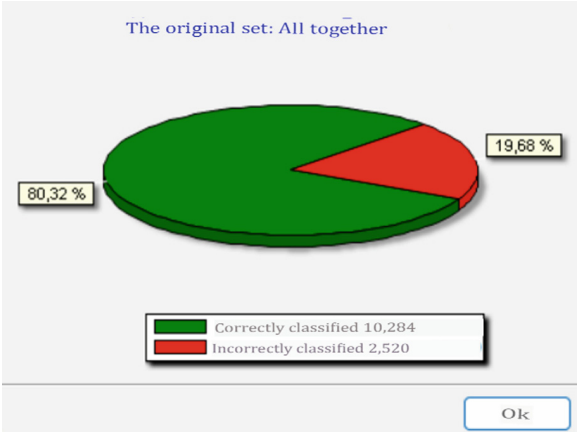


Fig. 28. Contiguity table

Next, the Kohonen map shown in Fig. 24 was constructed (Fig. 29).

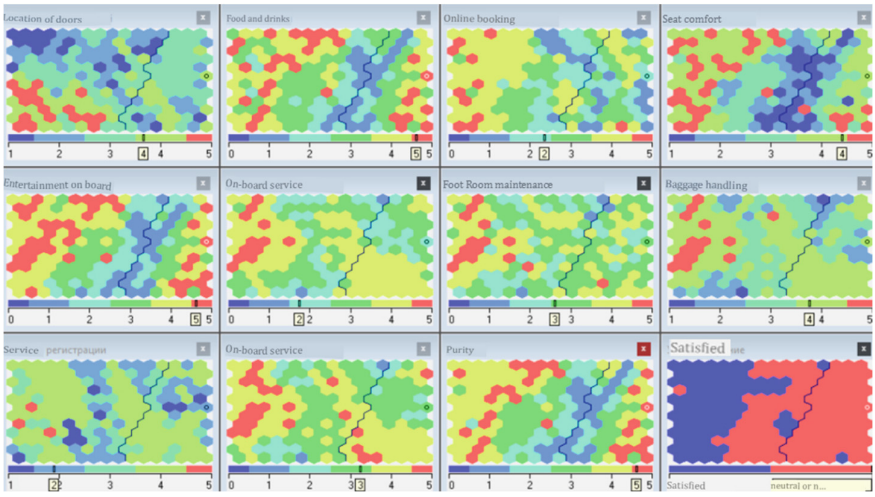


Fig. 29. Kohonen map

The total results of the two subgroups were almost identical, suggesting that gender did not play much of a role in this data set.

4.3 Type of Travel

A subset of business travel was considered first. At the beginning, the procedure of decision tree construction was carried out.

Target Attribute: Satisfaction

N	Number	Attribute	Significance, %
1	12	Online booking	44,879
2	7	Wi-Fi	17,910
3	14	Entertainment on board	7,172
4	13	Seat comfort	5,604
5	4	Customer Type	5,493
6	17	Baggage handling	3,551
7	5	Years	3,057
8	10	Location of doors	2,462
9	2	id	2,247
10	19	On-board service	1,776
11	6	Class	1,475
12	1	COL1	1,225
13	16	Foot Room maintenance	0,913
14	18	Registration Service	0,900
15	8	Convenient departure time/approx...	0,574
16	15	On-board service	0,316
17	20	Purity	0,282
18	3	Paul	0,105
19	11	Food and drinks	0,034

Fig. 30. Significance of attributes

Figure 30 shows a table showing the significance of the attributes. After that a neural network was built, with output data in the form of passenger satisfaction, Fig. 31 shows the contiguity table, which shows the good learnability of the neural network based on this data [44–49].

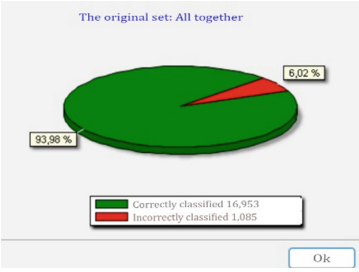


Fig. 31. Contiguity table

Clustering was then performed and the results are shown in Figs. 32, 33.

Fig. 32. Cluster Profiles

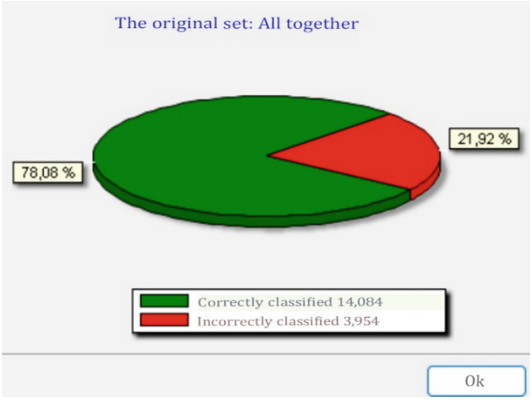


Fig. 33. Contiguity table

Next, the Kohonen map shown in Figs. 34 was constructed.

After that, a subset of personal travel was considered. At the beginning, the procedure of decision tree construction was carried out.

Figure 35 shows a table showing the significance of the attributes.

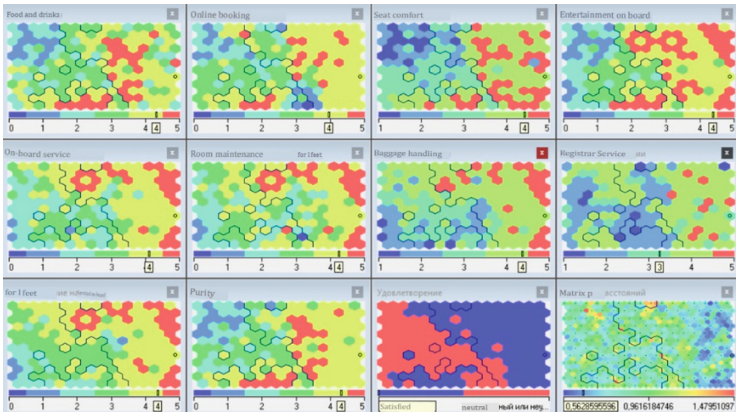


Fig. 34. Kohonen map

Target Attribute: Satisfaction			
№	Number	Attribute	Significance, %
1	8	Wi-Fi	99,609
2	2	id	0,267
3	1	COL1	0,124
4	17	Foot Room maintenance	0,000
5	18	Baggage handling	0,000
6	14	Seat comfort	0,000
7	15	Entertainment on board	0,000
8	16	On-board service	0,000
9	22	Average rating	0,000
10	23	Departure delay in minutes	0,000
11	21	Purity	0,000
12	19	Registration Service	0,000
13	20	On-board service	0,000
14	6	Class	0,000
15	7	Flight distance	0,000
16	5	Years	0,000
17	3	Pauli	0,000
18	4	Customer Type	0,000
19	12	Food and drinks	0,000
20	13	Online booking	0,000
21	11	Location of doors	0,000
22	9	Convenient departure time/approx...	0,000
23	10	Ease of online booking	0,000

Fig. 35. Significance of attributes

After that the neural network was built, with output data in the form of passenger satisfaction, Fig. 36 shows the contiguity table, which shows the good learning ability of the neural network based on this data.

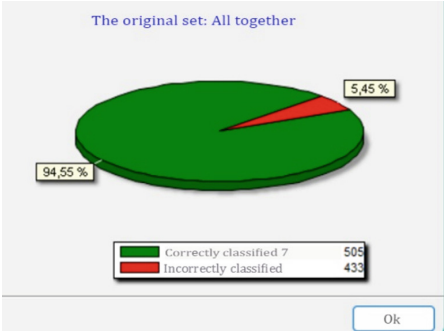


Fig. 36. Contiguity table

Clustering was then performed and the results are shown in Figs. 37, 38.

Fig. 37. Cluster Profiles

Next, the Kohonen map shown in Fig. 34 was constructed (Fig. 39).

Total, the significance of the factors of one subgroup is significantly different from the significance of the factors of the other subgroup. All other results showed an identical result.

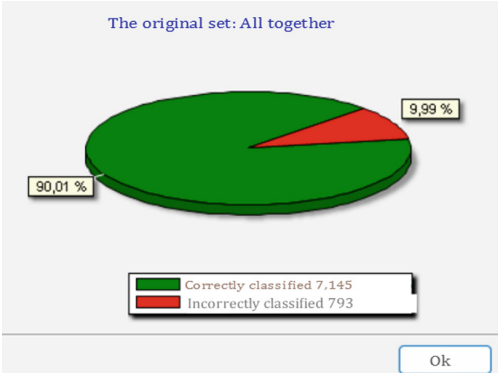


Fig. 38. Contiguity table

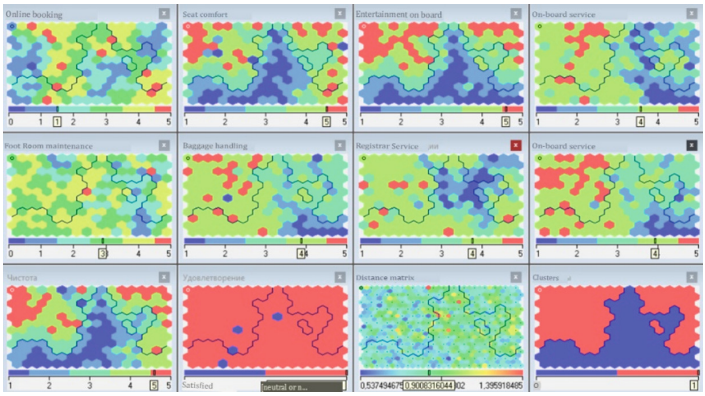


Fig. 39. Kohonen map

5 Conclusion

Based on the text presented, a dataset containing information on airline passengers' evaluations of various criteria was analyzed. The study identified several important factors affecting passenger satisfaction, such as baggage handling, seat comfort, cleanliness on board and availability of in-flight entertainment. The correlation analysis performed revealed strong relationships between different attributes, for example between baggage handling and on-board service or between cleanliness and food and beverage ratings. Neural networks were applied to predict satisfaction levels. Some models showed good learning and classification performance, while others had limited accuracy. Additionally, the dataset was categorized into semantic groups based on the class, gender and travel type of passengers. This allowed additional dependencies to be identified and more accurate decisions to be made.

Thus, the data analysis conducted allows airlines to better understand the factors affecting passenger satisfaction and take actions to improve the quality of service provided [49–51].

References

1. Gary, W.: Miller Chapter eight - Data science and the exposome, pp. 181–209 (2020)
2. Nagar, D., Pannerselvam, K., Ramu, P.: A novel data-driven visualization of *n*-dimensional feasible region using interpretable self-organizing maps (iSOM), vol. 155, pp. 398–412 (2022)
3. Tang, W., Li, W.: Frictional pressure drop during flow boiling in micro-fin tubes: a new general correlation, vol. 159, p. 120049 (2020)
4. Liu, Y., Jiang, Y., Hou, T., Liu, F.: A new robust fuzzy clustering validity index for imbalanced data sets, vol. 547, pp. 579–591 (2021)
5. Li, F., Zhang, X., Zhang, X., Du, C., Xu, Y., Tian, Y.-C.: Cost-sensitive and hybrid-attribute measure multi-decision tree over imbalanced data sets vol. 422, pp. 242–256 (2018)
6. Menzies, T., Kocagüneli, E., Minku, L., Peters, F., Turhan, B.: Chapter 6 - Rule #4: Data Science is Cyclic, pp. 35–38 (2015)
7. Comparison of Data Science Algorithms 2019, Pages 523–529
8. Zhu, C., Mei, C., Zhou, R.: Weight-based label-unknown multi-view data set generation approach, vol. 146, pp. 1–12 (2019)
9. Griffiths, G.W., Płociniczak, Ł., Schiesser, W.E.: Analysis of cornea curvature using radial basis functions – Part II: fitting to data-set, vol. 77, pp. 285–296 (2016)
10. Mariño, L.M.P., de Carvalho, A.T.F.: Vector batch SOM algorithms for multi-view dissimilarity data. Knowl. Based Syst. **258**, 109994 (2022). https://doi.org/10.1016/j.knosys.2022.109994
11. Mariño, L.M.P., de Carvalho, A.T.F.: Two weighted c-medoids batch SOM algorithms for dissimilarity data. Inform. Sci. **607**, 603–619 (2022). https://doi.org/10.1016/j.ins.2022.06.019
12. He, S.-F., Zhou, Q., Wang, F.: Local wavelet packet decomposition of soil hyperspectral for SOM estimation, vol. 125, p. 104285 (2022)
13. Zheng, Q., et al.: Multi-stage design space reduction technology based on SOM and rough sets, and its application to hull form optimization, vol. 213, part C, p. 119229 (2023)
14. Kang, H., Lee, K.S., Lee, H.Y., Chung, M.J., Yi, C.A., Kim, T.S.: CT Findings of Influenza A (H1N1) pneumonia in adults: pattern analysis and prognostic correlation. Chest **140**(4), 758A (2011). https://doi.org/10.1378/chest.1114485
15. Rubio-Rivas, M., Corbella, X.: Clinical phenotypes and prediction of chronicity in sarcoidosis using cluster analysis in a prospective cohort of 694 patients. Eur. J. Internal Med. **77**, 59–65 (2020). https://doi.org/10.1016/j.ejim.2020.04.024
16. Barchitta, M., et al.: Cluster analysis identifies patients at risk of catheter-associated urinary tract infections in intensive care units: findings from the SPIN-UTI Network, vol. 107, pp. 57–63 (2021)
17. Wang, R., Fung, B.C.M., Zhu, Y.: Heterogeneous data release for cluster analysis with differential privacy. Knowl. Based Syst. **201–202**, 106047 (2020). https://doi.org/10.1016/j.knosys.2020.106047
18. Carollo, A., Capizzi, P., Martorana, R.: Joint interpretation of seismic refraction tomography and electrical resistivity tomography by cluster analysis to detect buried cavities
19. Bosikov, I.I., et al.: Modeling and complex analysis of the topology parameters of ventilation networks when ensuring fire safety while developing coal and gas deposits. Fire **6**(3), 95 (2023)
20. Mikhalev A.S., et al.: The orb-weaving spider algorithm for training of recurrent neural networks. Symmetry **14**(10), 2036 (2022)
21. Moiseeva, K., et al.: The impact of coal generation on the ecology of city areas. In: 2023 22nd international symposium INFOTEH-JAHORINA (INFOTEH). IEEE, pp. 1–6 (2023)

22. Kukartsev, V., et al.: Analysis of Data in solving the problem of reducing the accident rate through the use of special means on public roads. In: 2022 IEEE international IOT, electronics and mechatronics conference (IEMTRONICS). IEEE, pp. 1–4 (2022)
23. Kireev, T., et al.: Analysis of the influence of factors on flight delays in the united states using the construction of a mathematical model and regression analysis. In: 2022 IEEE international IOT, electronics and mechatronics conference (IEMTRONICS). IEEE, pp. 1–5 (2022)
24. Kukartsev, V., et al.: Prototype technology decision support system for the EBW process. In: Silhavy, R., Silhavy, P., Prokopova, Z. (eds) Software engineering application in systems design. CoMeSySo 2022. LNNS, vol. 596. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-21435-6_39
25. Kukartsev, V., et al.: Methods and tools for developing an organization development strategy. In: 2022 IEEE international IOT, electronics and mechatronics conference (IEMTRONICS). IEEE, pp. 1–8 (2022)
26. Malozyomov, B.V.: Improvement of hybrid electrode material synthesis for energy accumulators based on carbon nanotubes and porous structures. Micromachines **14**(7), 1288 (2023)
27. Gutarevich, V.O., et al.: Reducing oscillations in suspension of mine monorail track. Appl. Sci. **13**(8), 4671 (2023)
28. Malozyomov, B.V., et al.: Overview of methods for enhanced oil recovery from conventional and unconventional reservoirs. Energies **16**(13), 4907 (2023)
29. Strateichuk, D.M., et al.: Morphological features of polycrystalline CdS1− xSex films obtained by screen-printing method. Crystals **13**(5), pp. 825 (2023)
30. Malozyomov, B.V., et al.: Study of supercapacitors built in the start-up system of the main diesel locomotive. Energies, **16**(9), 3909 (2023)
31. Malozyomov, B.V., et al.: Substantiation of drilling parameters for undermined drainage boreholes for increasing methane production from unconventional coal-gas collectors. Energies **16**(11), 4276 (2023)
32. Masich, I.S., Tyncheko, V.S., Nelyub, V.A., Bukhtoyarov, V.V., Kurashkin, S.O., Borodulin, A.S.: Paired patterns in logical analysis of data for decision support in recognition. Computation **10**(10), 185 (2022)
33. Masich, I.S., et al.: Prediction of critical filling of a storage area network by machine learning methods. Electronics **11**(24), 4150 (2022)
34. Barantsov, I.A., et al.: Classification of acoustic influences registered with phase-sensitive OTDR using pattern recognition methods. Sensors **23**(2), 582 (2023)
35. Bukhtoyarov, V.V., et al.: A study on a probabilistic method for designing artificial neural networks for the formation of intelligent technology assemblies with high variability. Electronics **12**(1), 215 (2023)
36. Rassokhin, A., Ponomarev, A., Karlina, A.: Nanostructured high-performance concretes based on low-strength aggregates. Magaz. Civil Eng. **110**(2), 11015 (2022)
37. Rassokhin, A., et al.: Different types of basalt fibers for disperse reinforcing of fine-grained concrete. Magaz. Civil Eng. **109**(1), 10913 (2022)
38. Shutaleva, A., et al.: Migration potential of students and development of human capital. Educ. Sci. **12**(5), 324 (2022)
39. Efremenkov, E.A., et al.: Research on the possibility of lowering the manufacturing accuracy of cycloid transmission wheels with intermediate rolling elements and a free cage. Appl. Sci. **12**(1), 5 (2021)
40. Shutaleva, A., et al. Environmental behavior of youth and sustainable development. Sustainability **14**(1), 250 (2021)
41. Repinskiy, O.D., et al.: Improving the competitiveness of Russian industry in the production of measuring and analytical equipment. J. Phys. Conf. Ser. IOP Publishing **1728**(1), 012032 (2021)

42. Balanovskiy, A.E., et al.: Determination of rail steel structural elements via the method of atomic force microscopy. CIS Iron Steel Rev. **23**, 86–91 (2022)
43. Kondrat'ev, V.V., et al.: Description of the complex of technical means of an automated control system for the technological process of thermal vortex enrichment. J. Phys. Conf. Ser. IOP Publishing **1661**(1), 012101 (2020)
44. Malozyomov, B.V., et al.: Improvement of hybrid electrode material synthesis for energy accumulators based on carbon nanotubes and porous structures. Micromachines **14**(7), 12888 (2023)
45. Potapenko, I., et al.: Analysis of the structure of germany's energy sector with self-organizing kohonen maps. In: Abramowicz, W., Auer, S., Stróżyna, M. (eds.) Business information systems workshops. BIS 2021. LNBIP, vol. 444. Springer, Cham (2021). https://doi.org/10.1007/978-3-031-04216-4_1
46. Borodulin, A.S., et al.: Using machine learning algorithms to solve data classification problems using multi-attribute dataset. E3S Web of Conferences. – EDP Sciences (2023)
47. Nelyub, V.A., et al.: Machine learning to identify key success indicators. E3S Web of Conferences. – EDP Sciences (2023)
48. Kukartsev, V.V., et al.: Using digital twins to create an inventory management system. E3S Web of Conferences. EDP Sciences (2023)
49. Gladkov, A.A., et al.: Development of an automation system for personnel monitoring and control of ordered products. E3S Web of Conferences. EDP Sciences (2023)
50. Kukartsev, V.V., et al.: Control system for personnel, fuel and boilers in the boiler house. E3S Web of Conferences. EDP Sciences (2023)
51. Kozlova, A.V., et al.: Finding dependencies in the corporate environment using data mining, E3S Web of Conferences. EDP Sciences (2023)
52. Lomazov, V.A., Lomazova, V.I., Miroshnichenko, I.V., Petrosov, D.A., Mironov, A.L.: Optimum planning of experimental research at the biogas plant. IOP Conf. Ser. Earth Environ. Sci. **659**(1), 012111 (2021). https://doi.org/10.1088/1755-1315/659/1/012111
53. Petrosov, D.A., Lomazov, V.A., Petrosova, N.V.: Model of an artificial neural network for solving the problem of controlling a genetic algorithm using the mathematical apparatus of the theory of petri nets (2021)
54. Petrosov, D.A., Lomazov, V.A., Klyuev, S.V., Mironov, A.L., Fomina, M.V.: Intellectual structural-parametric synthesis of large discrete systems with specified behavior. J. Eng. Appl. Sci. **13**(8), страницы 2177–2182 (2018)
55. Dmitriev, M.G., Lomazov, V.A.: Estimation of the linear convolution sensitivity of particular criteria during the expert determination of weight factors. Sci. Techn. Inform. Process. **41**(6), 400–403 (2014). https://doi.org/10.3103/S0147688214060033
56. Dmitriev M.G., Lomazov V.A.: Sensitivity of linear convolution from expert judgments, Procedia Computer Science. In: 2nd international conference on information technology and quantitative management, ITQM 2014, pp. 802–806 (2014)
57. Lomazov, V.A., Lomazov, A.V., Ivashchuk, O.A., Akupiyan, O.S., Nesterova, E.V.: Intellectual support for the analysis of the implementation of innovative development programs of the regional agro-industrial cluster. IOP Conf. Series Earth Environ. Sci. **839**(2), 022068 (2021). https://doi.org/10.1088/1755-1315/839/2/022068

Author Index

A
Ahmad, Anassuzastri 146
Akimenko, Tatiana 336, 343
Al Hamami, O. H. 239
Al Momani, Naim Mahmoud 91
Al-Ali, H. T. A. 112
Alekseev, Vladimir V. 239
Ali, Sadaquat 137
Alshahrani, Shadia Hamoud 91, 418
Andreeva, Maria 226
Anisimova, Tatyana 128
Ascencio, Joel 91, 418
Avsievich, Vladimir 397

B
Bakar, Abu 146, 165
Barot, Tomas 203, 210
Bayrak, L. A. 404
Behmaneshpour, Rafe 77
Belikov, Andrey 397
Bello, Abubakar 183
Bobrovs, Vjaceslavs 373
Bogomolov, Alexey 336, 343
Borodulin 434
Budinská, Ivana 77
Burden, Kapri 27
Bystrov, N. A. 404

C
Chebotarev, Andrey L. 245
Chertykovtseva, Natalia 397

D
Davzit, Ivan P. 245
Dineva, Vanya 194
Dmitrieva, Stepanida 226
Dubinina, Irina 428

E
El Amrani, Younes 91
Erkenova, J. I. 404, 411
Evdokarova, Tuara 226

F
Farid, Farnaz 183
Filipskaya, A. V. 411
Fomina, O. V. 239
Formánek, Tomáš 11
Frolov, Alexander 293
Frolova, O. A. 404

G
Gastelu, Angel Quiñonez 380
Gatial, Emil 77
Golosovskiy, Mikhail 336
Gunawan, P. H. 49
Gurov, Evgeny S. 245

H
Hanuranto, A. T. 49
Hasanah, Uswatun 146, 165
Hassankhani Dolatabadi, Sepideh 77
Hoang, Phuong Nguyen 137
Huda, Miftachul 146, 165

I
Ikonnikov, Oleg 128
Indwiarti 49
Irgasheva, Shokhida 418
Ismail, Mohamad Hazli 165
Ivanova, Saiyna 226
Ivaschenko, Anton 397, 428

J
Jahan, Nusrat 183
Joe, Inwhee 259

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
R. Silhavy and P. Silhavy (Eds.): CoMeSySo 2023, LNNS 910, pp. 459–460, 2024.
https://doi.org/10.1007/978-3-031-53552-9

K
Kalmuratov, Bekbosin 137
Kamaldinova, Zulfiya 428
Kappusheva, I. Sh. 404, 411
Kapustina, Svetlana 128
Karabtsev, Sergei N. 245
Karnaukhov, Andrey 137
Kenyeres, Martin 1
Khakimzyanov, Ruslan 137
Kim, HyunJoo 259
Kislov, Denis 373
Kleshko, I. I. 434
Klyagin, Anton 293
Kochetova, Natalya 293
Korabelshchikova, Svetlana 278, 286
Kosogorov, D. E. 352
Kotov, Roman M. 245
Krastev, Georgi 194
Krylatov, Alexander 100
Kubalcik, Marek 210
Kurbatov, Sergey S. 217
Kuzmich, Roman 128

L
Lakomov, D. V. 239
Laouar, M. A. 112
Larkin, Eugene 336, 343
Lavrinenko, E. D. 404
Losev, Aleksey 91
Lule, Emmanuel 61

M
Malele, Vusumuzi 37
Manrique, Mercedes Vega 380
Mapayi, Temitope 37
Marchuk, V. I. 112
Meghanathan, Natarajan 27, 363
Melnikov, Boris 278, 286
Mikeka, Chomora 61
Mirzagitova, Alsu 418
Mojžiš, Ján 1
Mphasane, Kagiso 37
Mujahidin, 146, 165
Mukanyiligira, Didacienne 61
Musdalifah, Parworth 61
Musolin, Mukhamad Hadi 146, 165

N
Nanay, F. A. 411
Nelyub, V. A. 434
Ngenzi, Alexander 61
Nikolaychuk, O. A. 352

P
Palupi, I. 49
Peña, Lidia Segura 380
Pestova, Yu. V. 352
Pupkov, Alexander 128

R
Raevskaya, Anastasiya 100
Robinson, Miah 27
Rohmawati, A. A. 49
Rudolf, Ladislav 210
Rukosueva 434

S
Samohleb, A. A. 112
Seberova, Alena 203
Senapati, Biswaranjan 128, 418
Shestakov, Viacheslav 128
Shestakova, Ekaterina 226
Shirokov, Ilya 428
Sogrina, V. A. 411
Solovyev, Alexander V. 119
Soria, Juan J. 380
Stanovova, D. A. 411
Sultana, Razia 183

T
Tolkacheva, Marianna 226
Torres, Alessandro Chanco 380
Tsarev, Roman 91, 128, 137, 418
Turkova, Vera 397
Tynchenko, V. S. 434

V
Vaclavik, Marek 203
Vasilevskiy, Valery V. 330
Vinh, Dang Van 278, 286
Voinohovska, Valentina 194

Y
Yauri, Andi Muhammad 146, 165

Z
Zhigalov, Kirill 91
Zuhri, Muhammad 146, 165
Zyablitseva, Larisa 278, 286