Annals of Computer Science and Information Systems
Volume 31

Position Papers of the 17th Conference on Computer Science and Intelligence Systems

September 4–7, 2022. Sofia, Bulgaria

Maria Ganzha, Leszek Maciaszek, Marcin Paprzycki,
Dominik Ślęzak (eds.)

Annals of Computer Science and Information Systems, Volume 31

Position Papers of the 17th Conference on Computer Science and Intelligence Systems

Maria Ganzha, Leszek Maciaszek, Marcin Paprzycki,
Dominik Ślęzak (eds.)

POLSKIE TOWARZYSTWO INFORMATYCZNE
POLISH INFORMATION PROCESSING SOCIETY

2022, Warszawa, Polskie
Towarzystwo Informatyczne

Annals of Computer Science and Information Systems, Volume 31

Position Papers of the 17th Conference on Computer Science and Intelligence Systems

USB: ISBN 978-83-965897-3-6
WEB: ISBN 978-83-965897-2-9

ISSN 2300-5963
DOI 10.15439/978-83-965897-2-9

Ul. Solec 38/103
00-394 Warsaw
Poland

Contact: secretariat@fedcsis.org
`http://annals-csis.org/`

Cover photo:
Dominik Żyłowski,
Elbląg, Poland

Also in this series:

Volume 32: Communication Papers of the 17th Conference on Computer Science and Intelligence Systems, **ISBN WEB: 978-83-965897-4-3, ISBN USB: 978-83-965897-5-0**
Volume 30: Proceedings of the 17th Conference on Computer Science and Intelligence Systems, **ISBN WEB: 978-83-962423-9-6, ISBN USB: 978-83-965897-0-5**
Volume 29: Recent Advances in Business Analytics. Selected papers of the 2021 KNOWCON-NSAIS workshop on Business Analytics**ISBN WEB: 978-83-962423-7-2, ISBN USB: 978-83-962423-6-5**
Volume 28: Proceedings of the 2021 International Conference on Research in Management & Technovation, **ISBN WEB: 978-83-962423-4-1, ISBN USB: 978-83-962423-5-8**
Volume 27: Proceedings of the Sixth International Conference on Research in Intelligent and Computing in Engineering, **ISBN WEB: 978-83-962423-2-7, ISBN USB: 978-83-962423-3-4**
Volume 26: Position and Communication Papers of the 16th Conference on Computer Science and Intelligence Systems, **ISBN WEB: 978-83-959183-9-1, ISBN USB: 978-83-962423-0-3**
Volume 25: Proceedings of the 16th Conference on Computer Science and Intelligence Systems, **ISBN WEB 978-83-959183-6-0, ISBN USB 978-83-959183-7-7, ISBN ART 978-83-959183-8-4**
Volume 24: Proceedings of the International Conference on Research in Management & Technovation 2020, **ISBN WEB: 978-83-959183-5-3, ISBN USB: 978-83-959183-4-6**
Volume 23: Communication Papers of the 2020 Federated Conference on Computer Science and Information Systems, **ISBN WEB: 978-83-959183-2-2, ISBN USB: 978-83-959183-3-9**
Volume 22: Position Papers of the 2020 Federated Conference on Computer Science and Information Systems, **ISBN WEB: 978-83-959183-0-8, ISBN USB: 978-83-959183-1-5**
Volume 21: Proceedings of the 2020 Federated Conference on Computer Science and Information Systems, **ISBN WEB 978-83-955416-7-4, ISBN USB 978-83-955416-8-1, ISBN ART 978-83-955416-9-8**
Volume 20: Communication Papers of the 2019 Federated Conference on Computer Science and Information Systems, **ISBN WEB: 978-83-955416-3-6, ISBN USB: 978-83-955416-4-3**

DEAR Reader, it is our pleasure to present to you Position Papers of the 17th Conference on Computer Science and Intelligence Systems (FedCSIS'2022), in Sofia, Bulgaria, and in the hybrid mode.

Position papers comprise two categories of contributions – challenge papers and emerging research papers. *Challenge papers* propose and describe research challenges in theory or practice of computer science and information systems. Papers in this category are based on deep understanding of existing research or industrial problems. Based on such understanding and experience, they define new exciting research directions and show why these directions are crucial to the society at large. *Emerging research papers* present preliminary research results from work-in-progress based on sound scientific approach but presenting work not completely validated as yet. They describe precisely the research problem and its rationale. They also define the intended future work including the expected benefits from solution to the tackled problem. Subsequently, they may be more conceptual than experimental.

The main Conference Chair of FedCSIS 2022 was Stefka Fidanova, while Nina Dobrinkova acted as the Chair of the Organizing Committee. This year, FedCSIS was organized by the Polish Information Processing Society (Mazovia Chapter), IEEE Poland Section Computer Society Chapter, Systems Research Institute Polish Academy of Sciences, Warsaw University of Technology, Wrocław University of Economics and Institute of Information and Communication Technologies, Bulgarian Academy of Sciences.

FedCSIS 2022 was technically co-sponsored by: IEEE Bulgarian Section, IEEE Poland Section, IEEE Czechoslovakia Section Computer Society Chapter, IEEE Poland Section Systems, Man, and Cybernetics Society Chapter, IEEE Poland Section Computational Intelligence Society Chapter, IEEE Poland Section Control System Society Chapter, Committee of Computer Science of the Polish Academy of Sciences, Mazovia Cluster ICT, Poland and Bulgarian Section of SIAM.

Moreover, last year the FedCSIS conference series formed the strategic alliance with QED Software, a Polish software company developing AI-based technologies and acting as the technological co-founder in the AI-driven start-ups. This collaboration has been continued.

During FedCSIS 2022, the keynote lectures were delivered by:

- Krassimir Atanassov, Bulgarian Academy of Sciences, Sofia, Bulgaria: "*Remarks on Index Matrices",*
- Chris Cornelis, Ghent University, Department of Applied Mathematics, Computer Science and Statistics: "*Hybridization of Fuzzy Sets and Rough Sets: Achievements and Opportunities",*
- Ivan Lukovic, University of Belgrade, Faculty of Organizational Sciences, *"Organizational Capability for Information Management – Do We Feel a Big Data Crisis?",*
- Stefano Mariani on behalf of Franco Zambonelli (due to serious health issues encountered right before the conference), University of Modena e Reggio Emilia, Italy: "*Individual and Collective Self-development*".

Moreover, this year, two special guests delivered invited talks:

- Bogusław Cyganek, who was awarded the 2021 Wiley-IEEE Press Award, for his recent book *"Introduction to Programming with C++ for Engineers",*
- Andrzej Skowron, who delivered the talk: "*Rough Sets Turn 40: From Information Systems to Intelligent Systems*", which was devoted to the 40th anniversary of introduction, by late Professor Zdzisław Pawlak, of the theory of Rough Sets.

FedCSIS 2022 consisted of five Tracks and a special event for Doctoral School Students. Within each Track, topical Technical Sessions have been organized. Each Technical Session was split into fully on site and fully online sub-sessions. The on-site part of the conference took place in the facilities of the Crisis Management and Disaster Response Centre of Excellence in Sofia, Bulgaria.

Some of Technical Sessions have been associated with the FedCSIS conference series for many years, while some of them are relatively new. The role of the technical Sessions is to focus and enrich discussions on selected areas, pertinent to the general scope of each Track. The list of Tracks, and topical Technical Sessions organized within their scope, was as follows.

- **Track 1: Advanced Artificial Intelligence in Applications (17th Symposium AAIA'22)**
  - Artificial Intelligence for Next-Generation Diagnostic Imaging (1st Workshop AI4NextGenDI'22)
  - Artificial Intelligence in Machine Vision and Graphics (4th Workshop AIMaViG'22)
  - Personalization and Recommender Systems (1st Workshop PeRS'22)
  - Rough Sets: Theory and Applications (4th International Symposium RSTA'22)
  - Computational Optimization (15th International Workshop WCO'22)
- **Track 2: Computer Science & Systems (CSS'22)**
  - Computer Aspects of Numerical Algorithms (15th Workshop CANA'22)
  - Concurrency, Specification and Programming (30th International Symposium CS&P'22)
  - Multimedia Applications and Processing (15th International Symposium MMAP'22)
  - Scalable Computing (12th Workshop WSC'22)
- **Track 3: Network Systems and Applications (NSA'22)**
  - Complex Networks - Theory and Application (1st Workshop CN-TA'22)
  - Internet of Things – Enablers, Challenges and Applications (6th Workshop IoT-ECAW'22)
  - Cyber Security, Privacy, and Trust (3rd International Forum NEMESIS'22)
- **Track 4: Advances in Information Systems and Technology (AIST'22)**
  - Data Science in Health, Ecology and Commerce (4th Workshop DSH'22)
  - Information Systems Management (17th Conference ISM'22)

- Knowledge Acquisition and Management (28th Conference KAM'22)

- **Track 5: Software and System Engineering (S3E'22)**
  - Cyber-Physical Systems (9th International Workshop IWCPS'22)
  - Model Driven Approaches in System Development (7th Workshop MDASD'22)
  - Software Engineering (42nd IEEE Workshop SEW'22)
- **7th Doctoral Symposium on Recent Advances in Information Technology (DS-RAIT'22)**

The program of FedCSIS 2022 required a dedicated effort of many people. We would like to express our warmest gratitude to all Committee members, of each Track and each Technical Session, for their hard work in attracting and later refereeing 290 submissions.

We thank the authors of papers for their great contribution to the theory and practice of computing and intelligence systems. We are grateful to the invited speakers for sharing their knowledge and wisdom with the participants.

Last, but not least, we thank Stefka Fidanova and Nina Dobrinkova. It should be stressed that they made all the preparations to organize the conference in Bulgaria for three years in a row, while only in 2022 the conference actually happened there. They also worked with us diligently to adapt the conference formula to organize it in hybrid mode. We are very grateful for all your efforts!

We hope that you had an inspiring conference. We also hope to meet you again for the 18th Conference on Computer Science and Intelligence Systems (FedCSIS 2023) which will take place in Warsaw, Poland on September 17-20, 2023.

Co-Chairs of the FedCSIS Conference Series:
Maria Ganzha, *Warsaw University of Technology, Poland and Systems Research Institute Polish Academy of Sciences, Warsaw, Poland*
Leszek Maciaszek*, Macquarie University, Sydney, Australia*
Marcin Paprzycki, *Systems Research Institute Polish Academy of Sciences, Warsaw Poland and Management Academy, Warsaw, Poland*
Dominik Ślęzak, *Institute of Informatics, University of Warsaw, Poland*

Position Papers of the 17th Conference on Computer Science and Intelligence Systems

September 4–7, 2022. Sofia, Bulgaria

TABLE OF CONTENTS

17$^{\text{TH}}$ International Symposium on Advanced Artificial Intelligence in Applications

4$^{\text{TH}}$ International Workshop on Artificial Intelligence in Machine Vision and Graphics

1$^{\text{ST}}$ Workshop on Personalization and Recommender Systems

Network Systems and Applications

1ST WORKSHOP ON COMPLEX NETWORKS: THEORY AND APPLICATION

6TH WORKSHOP ON INTERNET OF THINGS - ENABLERS, CHALLENGES AND APPLICATIONS

3RD INTERNATIONAL FORUM ON CYBER SECURITY, PRIVACY AND TRUST

ADVANCES IN INFORMATION SYSTEMS AND TECHNOLOGIES

28TH CONFERENCE ON KNOWLEDGE ACQUISITION AND MANAGEMENT

SOFTWARE, SYSTEM AND SERVICE ENGINEERING

JOINT 42ND IEEE SOFTWARE ENGINEERING WORKSHOP AND 9TH INTERNATIONAL WORKSHOP ON CYBER-PHYSICAL SYSTEMS

17th International Symposium Advances in Artificial Intelligence and Applications

THIS track is a continuation of international AAIA symposiums, which have been held since 2006. It aims at establishing the synergy between technical sessions, which encompass wide range of aspects of AI. With its longest-tradition threads, such as WCO focusing on Computational Optimization, it is also open to new initiatives categorized with respect to both, the emerging AI-related methodologies and practical usage areas. Nowadays, AI is usually perceived as closely related to the data, therefore, this track's scope includes the elements of Machine Learning, Data Quality, Big Data, etc. However, the realm of AI is far richer and our ultimate goal is to show relationships between all of its subareas, emphasizing a cross-disciplinary nature of the research branches such as XAI, HCI, and many others.

AAIA'22 brings together scientists and practitioners to discuss their latest results and ideas in all areas of Artificial Intelligence. We hope that successful applications presented at AAIA'22 will be of interest to researchers who want to know about both theoretical advances and latest applied developments in AI.

Topics

Papers related to theories, methodologies, and applications in science and technology in the field of AI are especially solicited. Topics covering industrial applications and academic research are included, but not limited to:

- Decision Support
- Machine Learning
- Fuzzy Sets and Soft Computing
- Rough Sets and Approximate Reasoning
- Data Mining and Knowledge Discovery
- Data Modeling and Feature Engineering
- Data Integration and Information Fusion
- Hybrid and Hierarchical Intelligent Systems
- Neural Networks and Deep Learning
- Reinforcement Learning
- Bayesian Networks and Bayesian Reasoning
- Case-based Reasoning and Similarity
- Web Mining and Social Networks
- Business Intelligence and Online Analytics
- Robotics and Cyber-Physical Systems
- AI-centered Systems and Large-Scale Applications
- AI for Combinatorial Games, Video Games and Serious Games
- Evolutionary Algorithms and Evolutionary Computation
- Artificial Intelligence for Next-Generation Diagnostis Imaging (1st Workshop AI4NextGenDI'22)
- Artificial Intelligence for Patient Empowerment with Sensor Systems (1st Workshop AI4Empowerment'22)
- Artificial Intelligence in Machine Vision and Graphics (4th Workshop AIMaViG'22)
- Intelligent Ambient Assisted Living Systems (1st Workshop IntelligentAAL'22)
- Personalization and Recommender Systems (1st Workshop PeRS'22)
- Rough Sets: Theory and Applications (4th International Symposium RSTA'22)
- Computational Optimization (15th Workshop WCO'22)

Track Chairs

- **Zdravevski, Eftim,** Ss. Cyril and Methodius University, Macedonia
- **Szczuka, Marcin,** University of Warsaw, Poland
- **Matwin, Stan,** Dalhousie University, Canada

Program Chairs

- **Corizzo, Roberto,** American University, USA
- **Sosnowski, Łukasz,** Systems Research Institute, Polish Academy of Sciences, Poland
- **Świechowski, ,** Maciej, QED Software, Poland

Program Committee

- **Azad, Mohammad,** Jouf University, Saudi Arabia
- **Bellinger, Colin,** National Research Council of Canada – Ottawa, Canada
- **Bianchini, Monica,** Dipartimento di Ingegnegneria dell'Informazione, Università di Siena, Italy
- **Boukouvalas, Zois,** American University – Washington DC, USA
- **Calpe Maravilla, Javier,** University of Valencia, Spain
- **Chelly, Zaineb,** Université Paris-Saclay, UVSQ, DAVID, France
- **Colantonio, Sara,** ISTI-CNR, Italy
- **Corizzo, Roberto,** American University, USA
- **Cyganek, Bogusław,** AGH University of Science and Technology, Poland
- **Dey, Lipika,** TCS Innovation Lab Delhi, India
- **Durães, Dalila,** Universidade do Minho, Portugal
- **Filipe, Vitor,** UTAD, Portugal
- **Girardi, Rosario,** UFMA, Brasil
- **Goleva, Rossitza,** New Bulgarian University, Bulgaria
- **Hullam, Gabor,** Budapest University of Technology and Economics, Hungary

- **Hussain, Shahid,** Institute of Business Administration, Pakistan
- **Jakovljevic, Niksa,** University of Novi Sad, Faculty of Technical Sciences, Bulgaria
- **Jaromczyk, Jerzy W.,** University of Kentucky, USA
- **Kaczmarek, Katarzyna,** Instytut Badań Systemowych Polskiej Akademii Nauk, Poland
- **Kasprzak, Włodzimierz,** Warsaw University of Technology, Poland
- **Kwaśnicka, Halina,** Wroclaw University of Technology, Poland
- **Laskov, Lasko,** New Bulgarian University, Bulgaria
- **Lerga, Jonatan,** Rijeka Tehnical University, Croatia
- **Lingras, Pawan,** Saint Mary's University, Canada
- **Ljubić, Sandi,** University of Rijeka, Faculty of Engineering, Croatia
- **Matwin, Stan,** Dalhousie Unbiversity, Canada
- **Meneses, Claudio,** Universidad Católica del Norte, Chile
- **Mignone, Paolo,** Università degli studi di Bari, Italy
- **Mihajlov, Martin,** Jozef Stefan Institute, Slovenia
- **Moore, Neil,** University of Kentucky, Department of Computer Science, USA
- **Moshkov, Mikhail,** King Abdullah University of Science and Technology, Saudi Arabia
- **Mozgovoy, Maxim,** University of Aizu, Japan
- **Muñoz, Andrés,** Universidad de Cádiz, Spain
- **Myszkowski, Paweł,** Wroclaw University of Science and Technology, Poland
- **Noguera, Manuel,** University of Granada, Spain
- **Pataricza, András,** Budapest University of Technology and Economics, Hungary
- **Peters, Georg,** Munich University of Applied Sciences & Australian Catholic University, Germany
- **Petrovska, Biserka,** Goce Delcev University, North Macedonia
- **Pires, Ivan Miguel,** Instituto de Telecomunicações, Universidade da Beira Interior, and Universidade de Trás-os-Montes e Alto Douro, Portugal
- **Po, Laura,** Universitá di Modena e Reggio Emilia, Italy
- **Porta, Marco,** University of Pavia, Italy
- **Przybyła-Kasperek, Małgorzata,** Uniwersytet Śląski, Poland
- **Raghavan, Vijay,** University of Louisiana at Lafayette, USA
- **Ramanna, Sheela,** Department of Applied Computer Science, University of Winnipeg, Canada
- **Rauch, Jan,** University of Economics, Prague, Czech Republic
- **Schaefer, Gerald,** Loughborough University, United Kingdom
- **Sosnowski, Łukasz,** Systems Research Institute, Polish Academy of Sciences, Poland
- **Spinsante, Susanna,** Università Politecnica delle Marche, Italy
- **Stańczyk, Urszula,** Silesian University of Technology, Poland
- **Stoean, Catalin,** University of Craiova, Romania
- **Stojanov, Riste,** Faculty of Computer Science and Engineering, Skopje
- **Subbotin, Sergey,** Professor, Zaporozhye National Technical University, Ukraine
- **Szczuka, Marcin,** University of Warsaw, Poland
- **Szczęch, Izabela,** Poznan University of Technology, Poland
- **Trajkovik, Vladimir,** Ss. Cyril and Methodius University, Faculty of Computer Science and Engineering, North Macedonia
- **Turukalo, Tatjana Loncar,** FTN, University of Novi Sad, Bulgaria
- **Unland, Rainer,** University of Duisburg-Essen, ICB, Germany
- **Verstraete, Jörg,** Instytut Badań Systemowych Polskiej Akademii Nauk, Poland
- **Weber, Richard,** Universidad de Chile
- **Zakrzewska, Danuta,** Institute of Information Technology, Technical University of Lodz, Poland
- **Zdravevski, Eftim,** Ss. Cyril and Methodius University, North Macedonia
- **Zielosko, Beata,** Univeristy of Silesia, Institute of Computer Science, Poland
- **Świechowski, Maciej,** QED Software, Poland
- **Żurek, Dominik,** AGH, Poland

DOI: 10.15439/2022F264
ISSN 2300-5963 ACSIS, Vol. 31

Application of Random Sampling in the Concept-Dependent Granulation Method

Radosław Cybulski
University of Warmia and Mazury,
in Olsztyn
ul. Słoneczna 54, 10-710 Olsztyn, Poland
Email: radoslaw.cybulski@uwm.edu.pl

Piotr Artiemjew
University of Warmia and Mazury,
in Olsztyn
ul. Słoneczna 54, 10-710 Olsztyn, Poland
Email: artem@matman.uwm.edu.pl

***Abstract*—Professor Zadeh in his works proposed the idea of grouping similar objects on the basis of certain similarity measures, thus initiating the paradigm of granular computing. He made the assumption that similar objects may have similar decisions. This natural assumption, operates in other scientific methodologies, e.g. methods based on k nearest neighbours, in reasoning by analogy and in rough set theory. The above assumption implies the existence of grouped information nodes (granules) and has potential applications in reducing the size of decision systems. The hypothesis has guided,, the creation of granulation techniques based on the use of rough inclusions (introduced by Polkowski and Skowron) - according to the scheme proposed by Polkowski. In their work, the possibility of a large reduction of the size of decision systems while maintaining the classification efficiency was verified in experimental works.**

In this paper, we investigate the possibility of using random sampling in the approximation of decision systems - as part of dealing with Big Data sets. We use concept-dependent granulation as a reference approximation method. Experiments on selected real-world data have shown a common regularity that gives a hint on how to apply random sampling for fast and effective size reduction of decision systems.

I. Introduction

IN THIS paper, we employ a granulation technique derived from rough set theory [3]. More specifically, we applied the concept-dependent granulation technique to reduce the size of decision systems [6], a methodology derived from the method proposed by Polkowski in paper [4] and extended in the paper [6]. A comprehensive research in this context is conducted in the monograph [5]. A demonstration of the decision system approximation using the concept-dependent method - showing the use of granulation to reduce the size of decision systems - can be seen in Table I. In this Table, we see how the granulation process allows us to reduce the size of the training systems while retaining the internal knowledge from the original systems. For example, for a radius of 0.682 we have a reduction in the number of objects of almost 98 percent while maintaining the original efficiency. The effectiveness of granulation methods (according to Polkowski's scheme) has been verified in many contexts and works with basically every popular classifier from SVM [8], decision trees [10] to neural networks [9]. The methods have also found applications in the context of steganography [11], preprocessing before feeding data into neural networks [9], in ensemble models [7], in classification processes [12], for absorbing missing values [13], in localization of mobile robots under magnetically variable conditions [14]. Due to the computational complexity of our techniques, an area for exploration that has not yet been adequately explored is their use for with methods for dealing with Big Data. The use of random sampling is our starting point in this area. In this paper, we use examples of relatively small decision systems for a simple illustration of the techniques. The application to big data of our method is to sample up to the size of the data that can be recalculated in the assumed time. Our results tentatively verify this possibility. As a reference classifier, we chose the kNN method, which is not a dedicated choice for our method. Any other classification method adapted to the granular data could be used to verify our assumptions.

The rest of the publication consists of the following sections. In Section II, we have an introduction to the methodology used in the paper, a demonstration of the granulation method and an indication of the classifier. In Section III, we present experimental results divided into two parts. Initial results showing cross-sectional classification performance with different radii, and detailed results with random sampling for selected sensible (giving variable results) granulation radii. We summarise the work in Section IV, where we also present our future research plans.

II. Research Methodology

In this section we will introduce our reference granulation technique and the classifier used.

A. *Refence granulatiom method - concept-dependent granulation*

Let us illustrate the operation of the concept-dependent granulation technique with an example. The system that is being granulated was generated by the Toy Decision system generator tool [1], [2].

Let us define

$$g^{cd}_{r_{gran}}(u_i) = \{u_j \in U_{trn} : \frac{|IND(u_i, u_j)|}{|A|} \geq r_{gran}$$

TABLE I
EXAMPLE OF CLASSIFICATION USING TOY DATA - MUSHROOM DATA SET. THE RESULTS PRESENTED HERE ARE FOR TWO SUCCESSIVE GRANULATIONS, THE FIRST BEING $layer_1$ THE SECOND $layer_2$.

	$layer_1$		$layer_2$	
r_{gran}	acc	$GranSize$	acc	$GranSize$
0.364	0.887	5.4	0.886	2
0.409	0.884	9.4	0.884	2
0.455	0.891	15.6	0.89	2
0.5	0.915	20.2	0.894	2.8
0.545	0.947	33.8	0.903	4.8
0.591	0.966	40.8	0.887	8.6
0.636	0.983	44.2	0.905	11.2
0.682	0.994	43.8	0.946	15.8
0.727	0.995	48	0.977	21.6
0.773	0.996	58	0.992	27
0.818	1	94.2	0.996	41.6
0.864	1	200.4	1	82.8
0.909	1	504.8	1	226.6
0.955	1	1762.8	1	947.6
1	1	6499.2	1	6499.2

TABLE II
EXEMPLARY DECISION SYSTEM: IRIS-SHORT, 5 ATTRIBUTES, 15 OBJECTS

Day	$a1$	$a2$	$a3$	$a4$	$class$
u_1	4.6	3.1	1.5	0.2	$Iris-setosa$
u_2	4.9	3.1	1.5	0.1	$Iris-setosa$
u_3	5.1	3.3	1.7	0.5	$Iris-setosa$
u_4	4.4	3.0	1.3	0.2	$Iris-setosa$
u_5	5.0	3.6	1.4	0.2	$Iris-setosa$
u_6	6.0	3.4	4.5	1.6	$Iris-versicolor$
u_7	5.9	3.2	4.8	1.8	$Iris-versicolor$
u_8	5.5	2.4	3.8	1.1	$Iris-versicolor$
u_9	6.6	3.0	4.4	1.4	$Iris-versicolor$
u_{10}	5.5	2.6	4.4	1.2	$Iris-versicolor$
u_{11}	6.8	3.2	5.9	2.3	$Iris-virginica$
u_{12}	6.9	3.1	5.1	2.3	$Iris-virginica$
u_{13}	6.5	3.0	5.2	2.0	$Iris-virginica$
u_{14}	6.7	3.0	5.2	2.3	$Iris-virginica$
u_{15}	7.7	2.8	6.7	2.0	$Iris-virginica$

$$and\ d(u_i) = d(u_j)\}$$

$$IND(u_i, u_j) = \{a \in A; a(u_i) = a(u_j)\}$$

U_{trn} is the universe of training objects,

and $|X|$ is the cardinality of set

The sample concept-dependent granules with a 0.25 radius, derived from decision systems from Table II look as follows,

$$g^{cd}_{0.25}(u_1) = \{u_1, u_2, u_4, u_5, \}$$
$$g^{cd}_{0.25}(u_2) = \{u_1, u_2, \}$$
$$g^{cd}_{0.25}(u_3) = \{u_3, \}$$
$$g^{cd}_{0.25}(u_4) = \{u_1, u_4, u_5, \}$$
$$g^{cd}_{0.25}(u_5) = \{u_1, u_4, u_5, \}$$
$$g^{cd}_{0.25}(u_6) = \{u_6, \}$$
$$g^{cd}_{0.25}(u_7) = \{u_7, \}$$
$$g^{cd}_{0.25}(u_8) = \{u_8, u_{10}, \}$$
$$g^{cd}_{0.25}(u_9) = \{u_9, u_{10}, \}$$
$$g^{cd}_{0.25}(u_{10}) = \{u_8, u_9, u_{10}, \}$$
$$g^{cd}_{0.25}(u_{11}) = \{u_{11}, u_{12}, u_{14}, \}$$
$$g^{cd}_{0.25}(u_{12}) = \{u_{11}, u_{12}, u_{14}, \}$$
$$g^{cd}_{0.25}(u_{13}) = \{u_{13}, u_{14}, u_{15}, \}$$

TABLE III
PART1 - TRIANGULAR INDISCERNIBILITY MATRIX FOR CONCEPT-DEPENDENT GRANULE GENERATION ($i < j$), DERIVED FROM TABLE II
$c_{ij} = 1,\ if\ \frac{|IND(u_i,u_j)|}{|A|} \geq 0.25\ and\ d(u_i) = d(u_j),\ 0,\ otherwise.$

	u_1	u_2	u_3	u_4	u_5	u_6	u_7	u_8
u_1	1	1	0	1	1	0	0	0
u_2		1	0	0	0	0	0	0
u_3			1	0	0	0	0	0
u_4				1	1	0	0	0
u_5					1	0	0	0
u_6						1	0	0
u_7							1	0
u_8								1

TABLE IV
PART2 - TRIANGULAR INDISCERNIBILITY MATRIX FOR CONCEPT-DEPENDENT GRANULE GENERATION ($i < j$), DERIVED FROM TABLE II
$c_{ij} = 1,\ if\ \frac{|IND(u_i,u_j)|}{|A|} \geq 0.25\ and\ d(u_i) = d(u_j),\ 0,\ otherwise.$

	u_9	u_{10}	u_{11}	u_{12}	u_{13}	u_{14}	u_{15}
u_1	0	0	0	0	0	0	0
u_2	0	0	0	0	0	0	0
u_3	0	0	0	0	0	0	0
u_4	0	0	0	0	0	0	0
u_5	0	0	0	0	0	0	0
u_6	0	0	0	0	0	0	0
u_7	0	0	0	0	0	0	0
u_8	0	1	0	0	0	0	0
u_9	1	1	0	0	0	0	0
u_{10}		1	0	0	0	0	0
u_{11}			1	1	0	1	0
u_{12}				1	0	1	0
u_{13}					1	1	1
u_{14}						1	0
u_{15}							1

$$g^{cd}_{0.25}(u_{14}) = \{u_{11}, u_{12}, u_{13}, u_{14}, \}$$
$$g^{cd}_{0.25}(u_{15}) = \{u_{13}, u_{15}, \}$$

Random coverage of training systems is as follows,

$$Cover(U_{trn}) = \{g^{cd}_{0.25}(u_2), g^{cd}_{0.25}(u_3), g^{cd}_{0.25}(u_4), g^{cd}_{0.25}(u_6)$$
$$, g^{cd}_{0.25}(u_7), g^{cd}_{0.25}(u_{10}), g^{cd}_{0.25}(u_{11}), g^{cd}_{0.25}(u_{15}), \}$$

TABLE V
CONCEPT-DEPENDENT GRANULAR REFLECTION OF THE EXEMPLARY TRAINING SYSTEM FROM TABLE II, IN RADIUS 0.25, 5 ATTRIBUTES, 8 OBJECTS; MV IS MAJORITY VOTING PROCEDURE (THE MOST FREQUENT DESCRIPTORS CREATE A GRANULAR REFLECTION)

Day	$a1$	$a2$	$a3$	$a4$	$class$
$MV(g^{cd}_{0.25}(u_2))$	4.6	3.1	1.5	0.2	$Iris-setosa$
$MV(g^{cd}_{0.25}(u_3))$	5.1	3.3	1.7	0.5	$Iris-setosa$
$MV(g^{cd}_{0.25}(u_4))$	4.6	3.1	1.5	0.2	$Iris-setosa$
$MV(g^{cd}_{0.25}(u_6))$	6.0	3.4	4.5	1.6	$Iris-versicolor$
$MV(g^{cd}_{0.25}(u_7))$	5.9	3.2	4.8	1.8	$Iris-versicolor$
$MV(g^{cd}_{0.25}(u_{10}))$	5.5	2.4	4.4	1.1	$Iris-versicolor$
$MV(g^{cd}_{0.25}(u_{11}))$	6.8	3.2	5.9	2.3	$Iris-virginica$
$MV(g^{cd}_{0.25}(u_{15}))$	6.5	3.0	5.2	2.0	$Iris-virginica$

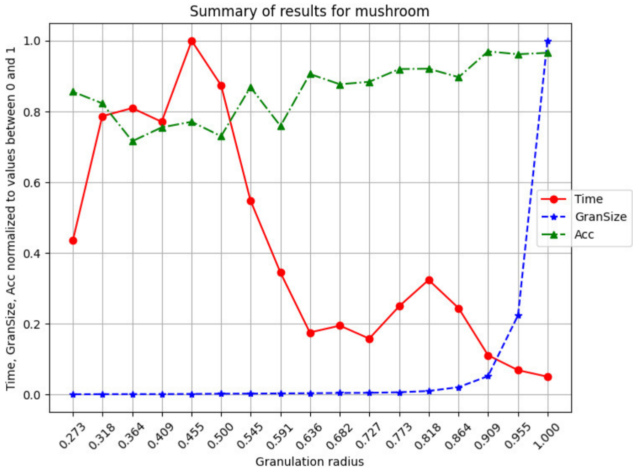


Fig. 1. Mushroom dataset - summary of results

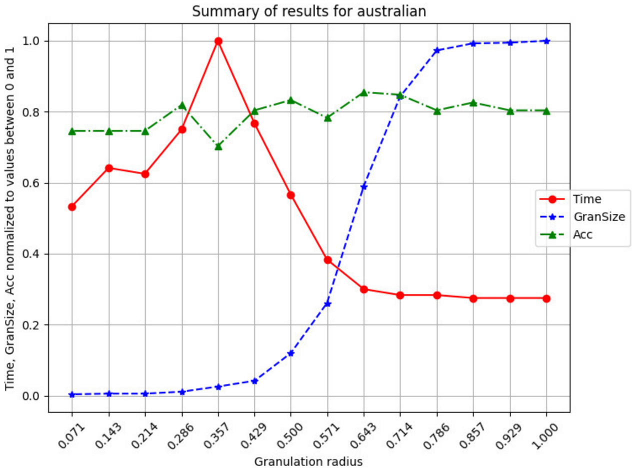


Fig. 2. Australian dataset - summary of results

The granulation process can be supported by using the indiscernibility matrix - see Tables III i IV. A granular reflection of the training system can be seen in Tab V.

B. Reference classifiers

1) Description of k-nearest neighbors algorithm: We use the kNN method from the Scikit-learn package as a reference classifier.

C. Random sampling

We use random sampling in our work on the basis of selecting a fixed percentage of objects - draw with return. In the results showing the use of this method on the x-axis we have the number of objects drawn. The parameters we present in our results are appropriately projected onto the interval [0,1]. We used standard normalization. The implementation was done in python language using standard libraries.

III. Experimental session

In the experimental part, we use three decision systems from the UCI repository [15], including Mushroom, Australian Credit and Heart Disease. In the kNN classifier we use k=1. For Mushroom and Heart we use the Euclidean metric, for Australian Hamming metric. For the initiation experiment, the data are split in a ratio of 0.8 to 0.2 and a cross-classification is performed for the granular systems created for the entire spectrum of granulation radii.

A. Reference results for concept- dependent granulation.

In the following, we will present a reference result for the granulation of the training system and the test classification - where the data are split in a ratio of 0.8 to 0.2. We take these results as a starting point for the analysis of the other results. Let us interpret the results the experiments, which are available in Figures 1, 2 and III-A.

When considering the approximation speed of decision systems, initial radii in the <0,0.5> range require training systems to be covered by a large number of granules which makes granulation slow. Once the threshold of 0.5 is exceeded, the granulation is already less time-consuming and the running time decreases. It is quite easy to find with this result areas where the radius of granulation is optimal, giving the result accuracy classification at a high level with a large reduction in the size of training systems. The optimal radius is a parameter that allows to achieve high classification accuracy (close to efficiency on full data). Our earlier discovery, the determination of optimal granulation radii by applying the layered granulation method, can also be used for this purpose - see [5]. When looking at the accuracy curve, we can see that the level of classification accuracy increases as the radii increase, this is due to the increase in the confidence of determining the classification parameters. At the same time the size of the granular decision systems increases, in the region of radius one, where we use the whole training system the classification level sometimes decreases because the noise existing in the data can be used for classification. We have shown previously that the granularity process for certain radii reduce the noise in the data - which increases the quality of the classification [5]. The last curve shows the percentage of granular systems in relation to the original training systems. It helps to determine in which area the granulation process should be completed. These overall granulation results are our starting point for research into the use of random sampling in tuning our granulation method. We show the results in the next section.

B. Concept-dependent granulation with random sampling.

The results, which are shown in Figures 4 to 21, demonstrate the interesting dependence of the granulation process on random sampling. By drawing a fixed percentage of objects from the original training set, we use a return draw. This causes that

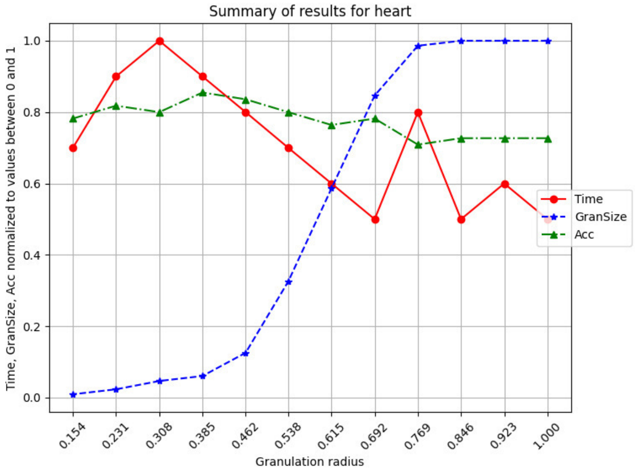


Fig. 3. Heart disease dataset - summary of results

in the granulation process some objects are absorbed. Hence, the percentage of granular systems in relation to original training systems starts to decrease with increasing random sample size.

First of all, for the individual radii, the decision-making system, regardless of the starting size of the random sample, has a similar final size. This can be observed by looking at the Gransize curve, where with increasing random sample the ratio of granular to pre-granular systems starts to decrease. At the same time, the classification accuracy shows a fairly high stability starting from radii in the region of 0.5. Which gives the conclusion that the use of random sampling significantly reduces the time required for the granulation process and allows the use of a strongly reduced random training sample. The level of reduction is individual to the specific data. In the decision systems studied, we observed that the classification accuracy is at a stable level with a reduction in the running time of the approximation of up to 80 percent over the full data.

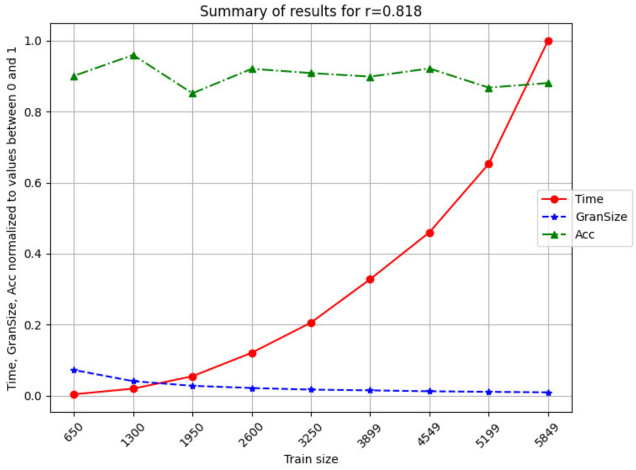


Fig. 4. Mushroom dataset - the result of random sampling for r=0.818; concept-dependent granulation

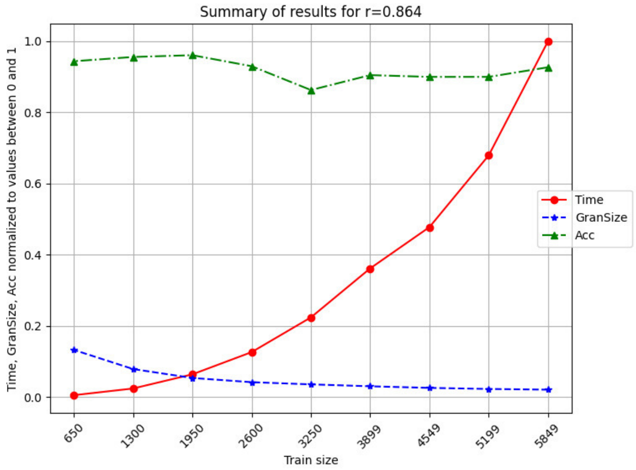


Fig. 5. Mushroom dataset - the result of random sampling for r=0.864; concept-dependent granulation

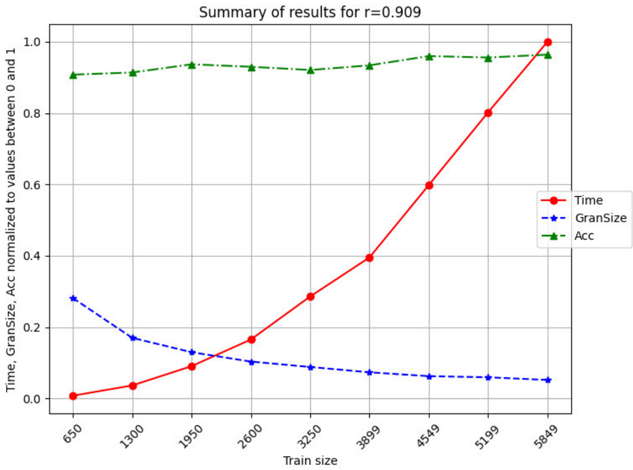


Fig. 6. Mushroom dataset - the result of random sampling for r=0.909; concept-dependent granulation

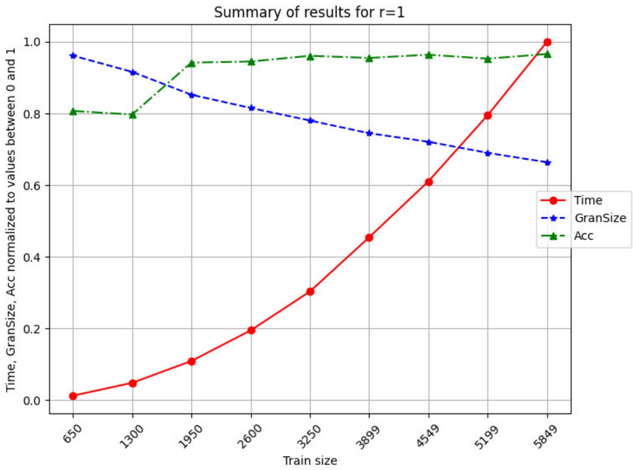


Fig. 8. Mushroom dataset - the result of random sampling for r=1.0; concept-dependent granulation

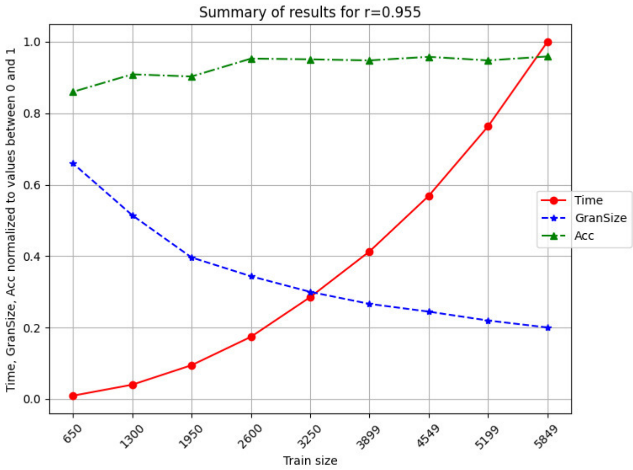


Fig. 7. Mushroom dataset - the result of random sampling for r=0.955; concept-dependent granulation

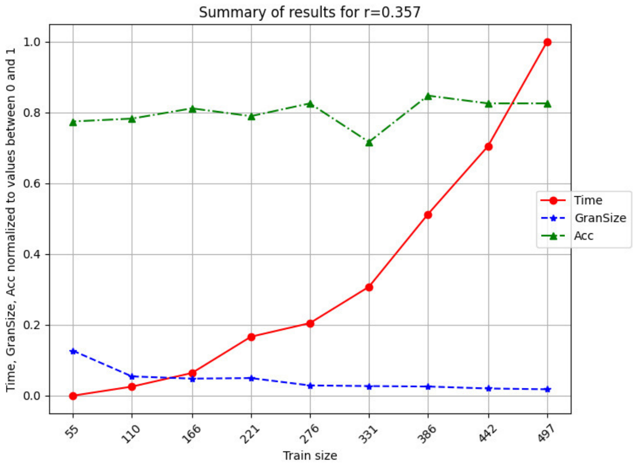


Fig. 9. Australian dataset - the result of random sampling for r=0.357; concept-dependent granulation

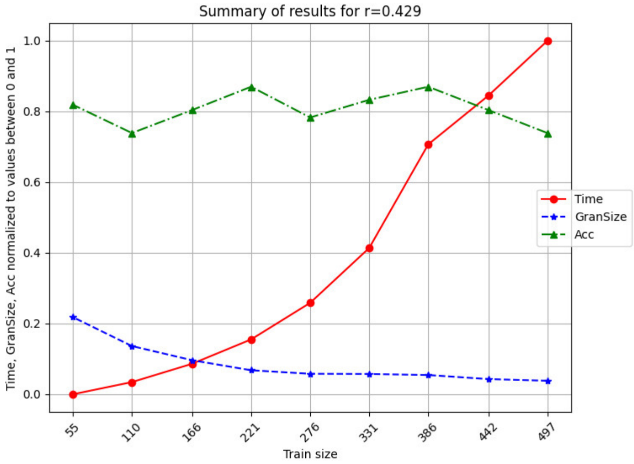


Fig. 10. Australian dataset - the result of random sampling for r=0.429; concept-dependent granulation

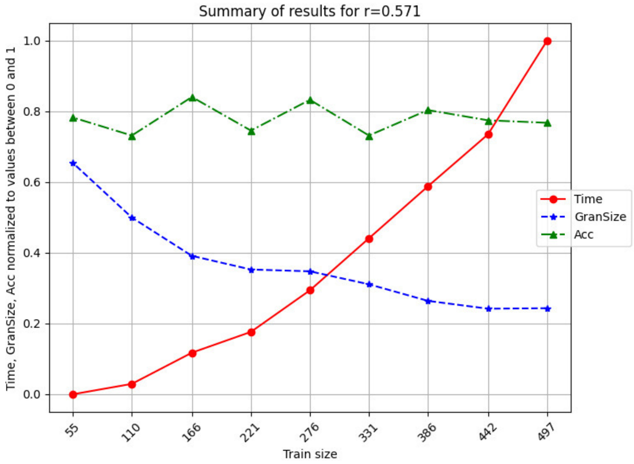


Fig. 12. Australian dataset - the result of random sampling for r=0.571; concept-depnedent granulation

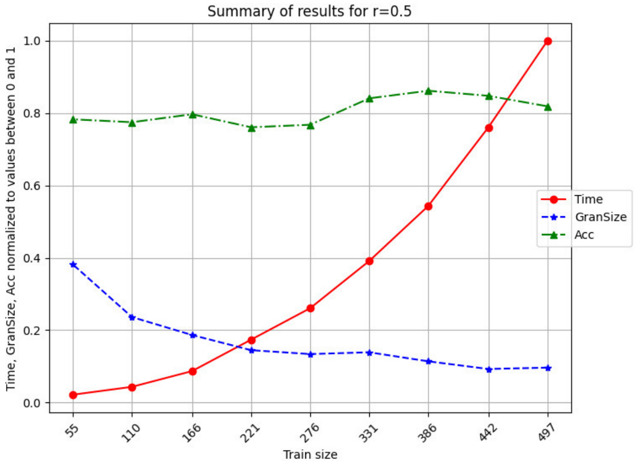


Fig. 11. Australian dataset - the result of random sampling for r=0.5; concept-dependent granulation

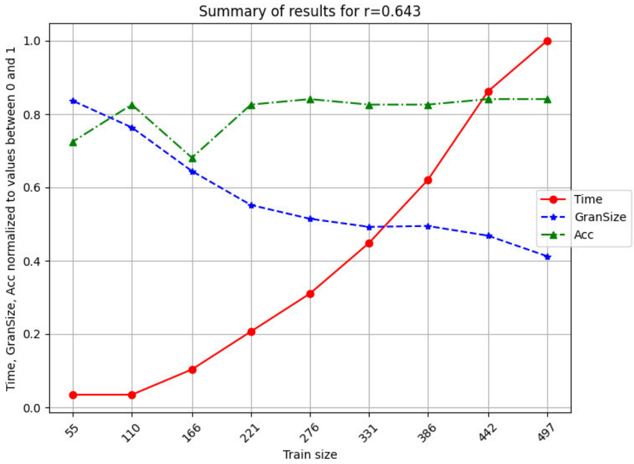


Fig. 13. Australian dataset - the result of random sampling for r=.0643; concept-dependent granulation

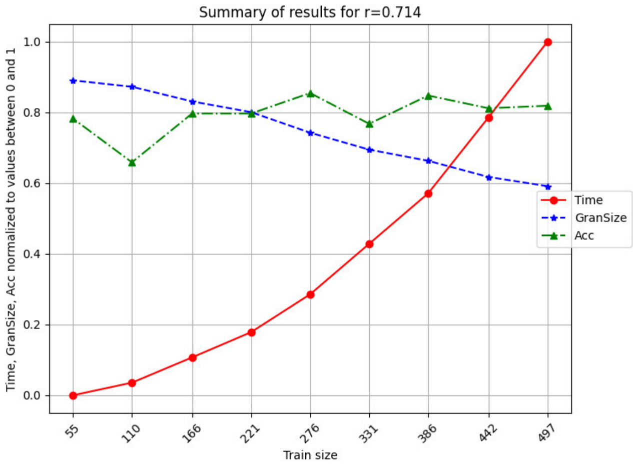


Fig. 14. Australian dataset - the result of random sampling for r=0.714; concept-dependent granulation

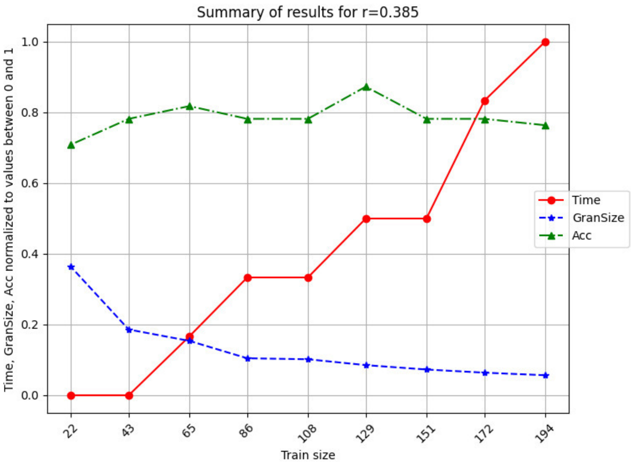


Fig. 16. Heart disease dataset - the result of random sampling for r=0.385; concept-dependent granulation

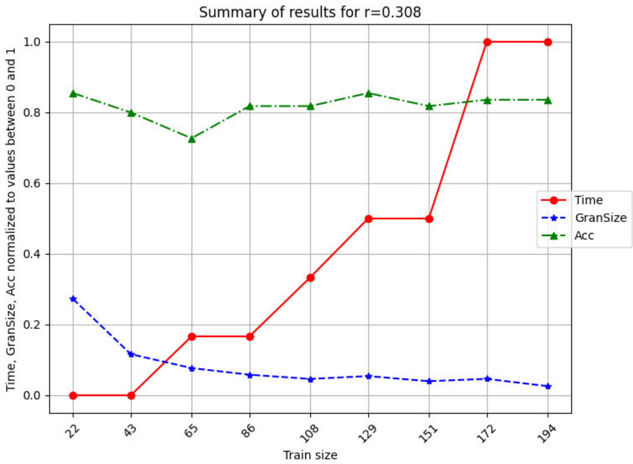


Fig. 15. Heart disease dataset - the result of random sampling for r=0.308; concept-dependent granulation

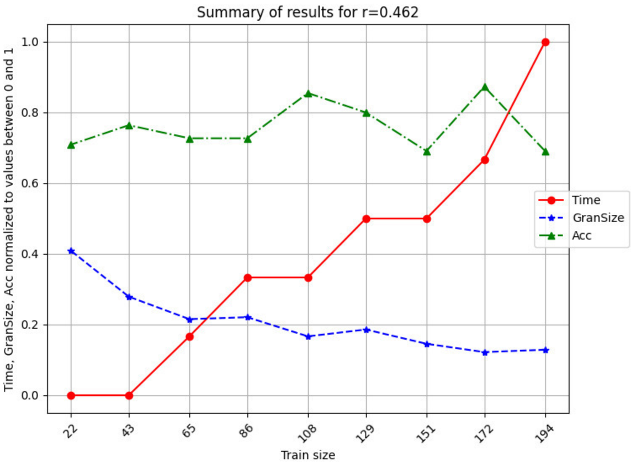


Fig. 17. Heart disease dataset - the result of random sampling for r=0.462; concept-dependent granulation

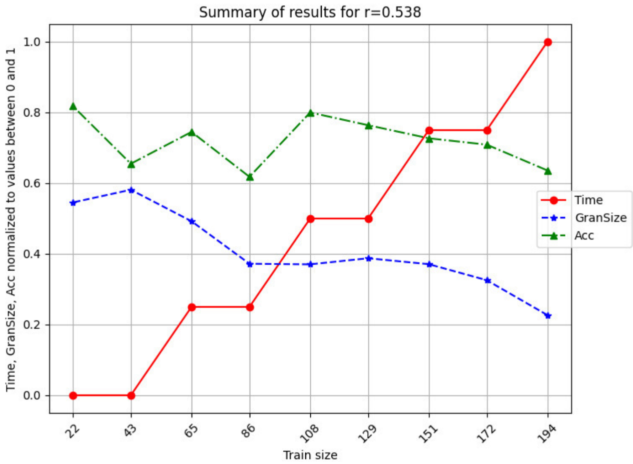


Fig. 18. Heart disease dataset - the result of random sampling for r=0.538; concept-dependent granulation

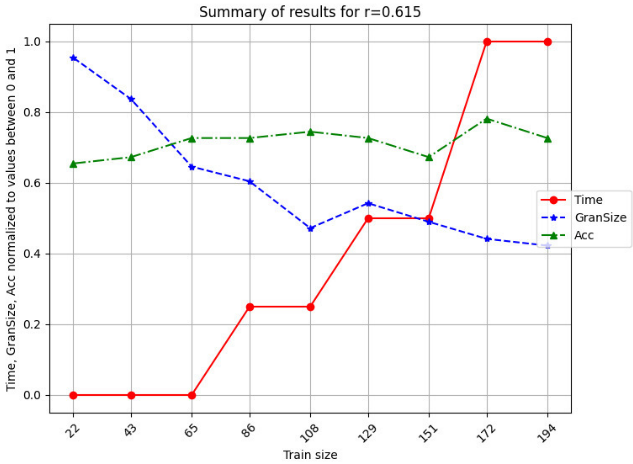


Fig. 19. Heart disease dataset - the result of random sampling for r=0.615; concept-dependent granulation

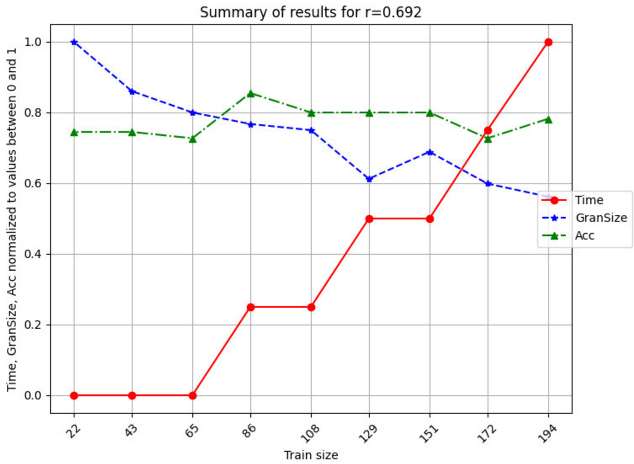


Fig. 20. Heart disease dataset - the result of random sampling for r=0.692; concept-dependent granulation

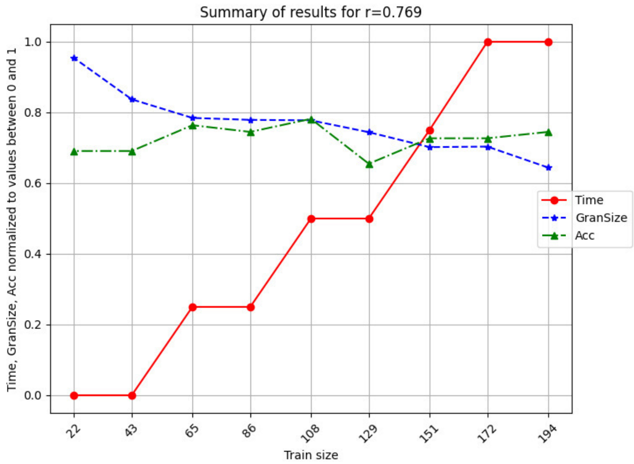


Fig. 21. Heart disease dataset - the result of random sampling for r=0.769; concept-dependent granulation

IV. Conclusion

In the current work, we made an interesting discovery - that random sampling works very well with the concept-dependent granulation method while maintaining the classification quality in reduced systems. For the decision systems studied, stable results, comparable to the performance of the original training systems - in terms of classification accuracy - are obtained with random sampling allowing us to reduce the running time of our method to as much as 20 percent of the original time (time is reduced by 80 percent). The current result was confirmed on three selected systems from the UCI repository, and represents an initial pilot study that opens new research horizons - using granulation methods based on approximate inclusions in the context of Big Data. An interesting observation is that the final granular systems for specific granulation radii (up to 0.5) have a similar size for individual random samples.

The subject of further research will be to look for a way to discover threshold values of random samples that give meaningful granulation results. In addition, we plan to explore in the context of granulation the whole range of possible techniques used for dealing with Big Data sets.

Acknowledgment

This work has been supported by the grant from Ministry of Science and Higher Education of the Republic of Poland under the project number 23.610.007-000

References

[1] Toy decision system generator, http://toyds.herokuapp.com/generator/v1/. Last accessed 12 Apr 2022

[2] Artiemjew, P.: (2022). Rough Inclusion Based Toy Decision Systems Generator For Presenting Data Mining Algorithms. Proceedings of the 3rd Polish Conference on Artificial Intelligence, April 25-27, 2022, Gdynia, Poland, 168-171.

[3] Pawlak, Z.: Rough sets. International Journal of Computer and Information Sciences 11, 341—356 (1982). https://doi.org/10.1007/BF01001956

[4] Polkowski, L.: A model of granular computing with applications, In: Proceedings of IEEE 2006 Conference on Granular Computing GrC06, pp. 9–16. IEEE Press, Atlanta, USA (2006)

[5] Polkowski, L., Artiemjew, P.: Granular Computing in Decision Approximation - An Application of Rough Mereology, In: Intelligent Systems Reference Library 77, Springer, ISBN 978-3-319-12879-5, pp. 1–422 (2015).

[6] Artiemjew, P.: Classifiers from Granulated Data Sets: Concept Dependent and Layered Granulation, In: Proceedings RSKD'07. The Workshops at ECML/PKDD'07, pp. 1–9., Warsaw Univ. Press, Warsaw (2007)

[7] Artiemjew, P., Ropiak, K.: 'A Novel Ensemble Model - The Random Granular Reflections', Fundamenta Informaticae, 1 Jan. 2021, vol. 179, no. 2, pp. 183-203, 2021(DOI: 10.3233/FI-2021-2020)

[8] J. Szypulski, P. Artiemjew: The Rough Granular Approach to Classifier Synthesis by Means of SVM, In: Proceedings of International Joint Conference on Rough Sets, IJCRS'15, pp. 256-263, Tianjin, China, Lecture Notes in Computer Science (LNCS), Springer, Heidelberg (2015)

[9] Ropiak, K.; Artiemjew, P. On a Hybridization of Deep Learning and Rough Set Based Granular Computing. Algorithms 2020, 13, 63.

[10] Ropiak K., Artiemjew P. (2020) Random Forests and Homogeneous Granulation. In: Lopata A., Butkiene R., Gudoniene D., Sukacke V. (eds) Information and Software Technologies. ICIST 2020. Communications in Computer and Information Science, vol 1283. Springer, Cham. https://doi.org/10.1007/978-3-030-59506-7_16 (2020)

[11] Artiemjew P., Kislak-Malinowska A. (2019) Using r-indiscernibility Relations to Hide the Presence of Information for the Least Significant Bit Steganography Technique. In: Damaševičius R., Vasiljeviene G. (eds) Information and Software Technologies. ICIST 2019. Communications in Computer and Information Science, vol 1078. Springer

[12] Artiemjew P.: Boosting effect for classifier based on simple granules of knowledge. In: Information Technology And Control (ITC) vol. 47(2), pp. 184-196 (2018)

[13] L. Polkowski, P. Artiemjew: Granular Computing: Classifiers and Missing Values, in Proceedings ICCI'07. 6th IEEE International Conference on Cognitive Informatics, IEEE Computer Society, Los Alamitos, CA, 2007, pp. 186-194.

[14] Artiemjew, P., Ropiak, K.: Robot localization in the magnetic unstable environment, 5th Workshop on Collaboration of Humans, Agents, Robots, Machines and Sensors (CHARMS 2019), The Third IEEE International Conference on Robotic Computing (IRC 2019), Naples, Italy

[15] UCI ML Repository, https://archive.ics.uci.edu/ml/index.php.

DOI: 10.15439/2022F279
ISSN 2300-5963 ACSIS, Vol. 31

Searching For Loops And Sound Samples With Feature Learning

Jan Jakubik
Wroclaw University of Science and Technology
Faculty of Information and Communication Technology
Department of Artificial Intelligence
jan.jakubik@pwr.edu.pl

***Abstract*—In this paper, we evaluate feature learning in the problem of retrieving subjectively interesting sounds from electronic music tracks. We describe an active learning system designed to find sounds categorized as samples or loops. These retrieval tasks originate from a broader R&D project, which concerns the use of machine learning for streamlining the creation of videogame content synchronized with soundtracks. The method is expected to function in the context of limited data availability, and as such cannot rely on supervised learning of what constitutes an "interesting sound". We apply an active learning procedure that allows us to find sound samples without predefined classes through user interaction, and evaluate the use of neural network feature extraction in the problem.**

***Index Terms*—music information retrieval, machine learning, signal processing**

I. Introduction

THE USE of machine learning methods in Music Information Retrieval (MIR) has developed significantly in the past decade thanks to the improvements in machine learning areas such as deep neural networks [1] and increased availability of big data. We now have large datasets available for problems such as genre recognition and auto-tagging [2], emotion recognition [3], and more specialized problems such as pitch tracking have seen massive improvements too [4]. The most limiting factor for many narrow MIR problems remains the lack of massive training datasets required to train well-performing deep models.

This issue becomes more problematic when we consider the limitations of existing datasets in practical applications. Many methods are continuously developed on existing well-defined problems, but when new practical demands arise, it is often hard to find appropriate data. The MIR problems described in this work were defined in cooperation with a business partner interested in streamlining the creation of music-synchronized videogame content. Such content relies on the ability to trigger in-game events, such as playing particle effects or animations, in sync with the audio. One of the functionalities the developer was interested in was a retrieval system in which the creator points to a single example of a particular sound in the context of an existing music track, and all other occurrences of that sound can be automatically marked.

The desired scenario creates an ambiguity impossible to resolve with just a single sample. As such, we have opted for an active learning solution in which the annotations are obtained through the interaction of the developer with the retrieval results. Our goal was to limit the effort of the user in finding other occurrences of the sound (ideally working perfectly with just a single occurrence). The system was also expected to perform in an open-set recognition scenario, where sound types cannot be pre-defined.

In this work, we focus on empirical evaluation of deep feature learning to the described retrieval scenario. We have previously published early results relying on the use of feature extraction techniques considered standard for music audio in this task [5]. We build upon the earlier work by extending the method through the use of deep feature learning and evaluate the result on an improved version of the dataset. The improved dataset contains more annotations and a clear distinction between two categories of repeating sounds typically found in electronic music - *loops* and *samples*.

II. Related Work

Retrieval of sound effects has been an active topic in MIR, and we can relate our problem to multiple existing ones. Sound event detection and classification [7] were considered supervised tasks in multiple contexts, including non-musical ones. These supervised tasks are usually defined as retrieval of specific, predefined sound classes from large-scale collections, which makes the supervised methods ill-fitting for our problem. However, zero-shot learning approaches which relate to the way our task is defined found success in sound effect classification [8] and could potentially be applied to ours. These rely on a pre-trained deep neural representation and transfer learning. Source separation [9] and onset detection [10] are related to our task in that we need to separate the sample and detect its times of appearance. There are onset detection datasets [11] for a limited range of sounds, usually drums, that could potentially serve as a benchmark for that type of sound only. However, our base assumption about the desired functionality of the system is that there are no limitations on the types of sounds users can mark as interesting.

Loop discovery has been considered in several papers, although it is not a very active research topic. In the area of music structure analysis, there is a concern with finding repeated patterns [12]. However, more relevant to our application, there exists a loop retrieval approach based on tensor decomposition introduced in [14]. The idea originates from [13] in which modeling of audio using predefined loops was attempted. The tensor decomposition approach improves upon that work by being able to decompose audio without loops being known in advance. It was later developed to create Unmixer [15], a publicly available system for fully unsupervised loop decomposition.

Active learning has been applied to MIR tasks on multiple fronts: genre recognition [16], mood recognition [17] and narrow tasks such as singing voice detection [18]. These works assume a scenario in which the goal is to maximize performance given a limited annotation budget. As such, they use a variety of metrics to select samples that are most beneficial to annotate. Metrics for sample selection can be broadly classified into two categories: based on uncertainty and based on correlation [19]. In the first case, the most uncertain samples or samples that would result in the largest model change are chosen. In the second case, samples are chosen to represent a significant subset of the data, e.g., each sample is representative of a particular cluster obtained in dataset clustering.

Unsupervised feature learning focuses on using large amounts of unannotated data to train general-purpose neural networks that can find use in downstream tasks with a small amount of training data. Early approaches to this problem usually utilize an encoder-decoder architecture [20] and the information bottleneck principle. A network that first encodes and then reconstructs the data implicitly creates a compact and robust lower-dimensional representation that leverages patterns within the unlabeled dataset. More recently, a lot of attention has been given to contrastive learning which leverages the power of data augmentation. A self-supervised network is trained by comparing data created from a single sample through different randomized augmentations, with the goal of creating a representation that is invariant to the augmentation. It has been shown that using the principle of contrastive learning alone is sufficient to train robust representations without labels for any supervised tasks [21].

III. Materials and Methods

Below we describe all data and methods used in the study. The dataset has been previously used in [5], but here it is developed further with the separation of two distinct types of sounds. The method of active retrieval and the feature representation it uses are descrived in subsections B-D.

A. Dataset Description

The dataset consists of 300 songs from the Creative Commons repository sampleswap.org. Audio files within the dataset are complete songs, ranging from 2 to 7.5 minutes in length. The songs have been selected from 4 musical genres (House, Dubstep, Drum&Bass and Downtempo) and annotated by three workers based on their subjective perceptions of interesting *sound samples* and *loops*. In the creation of electronic music, a *sound sample* is a pre-recorded sound that can be used for its interesting sonic qualities, while *loop* is a pattern that can seamlessly repeat, usually with both melodic and percussive components. Note that these aren't mutually exclusive, as any sound sample can be used within a loop, and any loop can be sampled. Our game developer partner was interested in retrieving both reused samples and actual loops, which lead to our attempt to develop a general method for any "standout" repeatable sounds, while still maintaining the distinction between both categories in the dataset.

When defining these concepts to the annotators, we asked them to consider *samples* to be audio effects and characteristic sounds that stand out against the musical background and can be heard at least twice in the same track, whereas *loops* were described as seamlessly repeating musical and rhythmic patterns. The annotations were created with 0.1-second precision. Within the *sample* category, annotators preferred short sounds: the majority of the sounds chosen were less than 3 seconds long. However, some persistent background sounds as long as 24 seconds were perceived and marked as a single sample of interest. *Loops* were longer on average, however, some loops as short as 2 seconds also occur in the dataset. Overall, there is a decent variety of what a potential user could understand as *samples* and *loops* represented within the dataset.

B. Active Learning Retrieval Approach

The desired system works as follows: given an audio file representation X, time of occurrence t_0 and duration d, the goal is to find a set of times $\{t_1, ..., t_n\}$ marking all other occurrences of this sound within the audio file. This search is performed according to Algorithm 1, which repeatedly polls the user with new retrieval results and then adds the responses to the growing set of positive samples P or negative samples N. The function $UserResponse(newsample)$ corresponds to the user giving a yes/no answer whether $newsample$ is a correct result. The algorithm is limited by $patience$, a parameter that represents the number of negative samples that can be returned before the user gives up on searching.

Algorithm 1 Active Retrieval Procedure

```
function RETRIEVE($X, t_0$, $d$, $patience$)
    $P \leftarrow \{t_0\}$
    $N \leftarrow \emptyset$
    while $|N| < patience$ do
        $newsample \leftarrow GetBestSamples(X, P, N, d)$
        if $UserResponse(newsample) = true$ then
            $P \leftarrow P \cup \{newsample\}$
        else
            $N \leftarrow N \cup \{newsample\}$
        end if
    end while
    return $P$
end function
```

The key issue in defining a method for solving this problem is the implementation of the function $GetBestSamples$ which uses some representation of the sound file X, the set of samples identified as positive so far P and the set of samples identified as negative so far N. This function should retrieve the most fitting candidate for a new positive sample, and return its time of occurrence. A natural choice for this function is nearest neighboor search in a feature space that represents the percieved similarity between sound excerpts well. In that case, the key element becomes the choice of the feature space.

C. Feature Representations of Audio

The baseline methods we present can be applied to multiple representations of audio, including a vector sequence obtained from a pre-trained deep learning model. In evaluation, we focus on the following vector sequence representations derived from the audio spectrogram:

1) Mel spectrogram: Mel spectrogram is an example of a spectral representation that takes the psychoacoustic properties of sound into account. Mel spectrogram transforms short frames of the signal into the frequency domain, using a logarithmically spaced Mel frequency spectrum. This corresponds to human perception of frequency better than the linearly spaced Short Time Fourier Transform. Mel spectrogram is not used directly as a feature representation (in preliminary tests, it achieved worse results than MFCC), but instead, serves as an input to a feature learning network described in subsection D.

2) MFCC: Mel-frequency Cepstral Coefficients are features commonly used in speech recognition that have found success in multiple MIR tasks. MFCC vectors capture timbral properties of sound well but lose precise frequency information.

D. Unsupervised Feature Learning

Our unsupervised feature learning setup combines autoencoding and contrastive learning losses. We have found that using only one of these was not sufficient, which will be elaborated on in Section IV. For the contrastive loss, we chose to base the network on the BYOL (Bootstrap Your Own Latent) approach [22], an evolution of the earlier SimCLR method [23]. For the autoencoding objective, we use a standard MSE reconstruction loss with no additional modifications.

1) Bootstrap Your Own Latent: The BYOL approach is an evolution of earlier contrastive learning methods, resulting from the observation that the previous methods took some unnecessary precautions from creating a loss with trivial, bad global optima. The loss $\mathcal{L}_{BYOL}$ for a pair of samples (x, x') created through data augmentation is written in a simplified form in Eq. 1:

$$\mathcal{L}_{BYOL}(x, x') = \|N(Pr(P(E(x)))) - N(P_f(E_f(x')))\|^2 \tag{1}$$

Two samples resulting from data augmentation: x and x', pass through four consecutive components: E denotes an encoder network, P denotes a projection network, Pr denotes a predictor network and N denotes vector normalization. The f subscript denotes the "frozen" version of components (i.e., not updated through gradient descent steps).

While network E is the feature extractor we are trying to obtain as the end goal of feature learning, other components exist to improve the training procedure. The projection layer P is optional but has been shown by the authors of the original SimCLR paper to improve the quality of trained representations in downstream tasks. The predictor network Pr helps prevent the collapse of training by making the architecture asymmetric. Unlike earlier contrastive learning methods, BYOL does not explicitly prevent a collapse to a bad global optimum in its loss (for example, if the encoder E outputs the same vector for any input, the MSE could be easily reduced to 0). However, the creators of the method have shown empirically that in a practical setting, with random initialization and gradient descent training of the N, P, and Pr components, the training procedure is not expected to collapse.

E, P and Pr networks are trained with gradient descent. E_f and P_f components are instead updated as an exponential running mean of respectively E and P. I.e., the parameters θ_f of a frozen network are updated based on the parameters θ of a respective unfrozen network, using Eq. 2 with a hyperparameter $\alpha \in (0, 1)$:

$$\theta_f = \alpha\theta_f + (1 - \alpha)\theta \tag{2}$$

2) Autoencoder Network: For autoencoder training objective $\mathcal{L}_{AE}$, we use the simplest possible formulation, as shown in Eq. 3:

$$\mathcal{L}_{AE}(x) = \|D(E(x)) - x\|^2 \tag{3}$$

The loss is calculated on an example x using encoder network E and decoder network D.

IV. Results

The results presented below are obtained through Algorithm 1 with $patience$ set to 5 and nearest neighboor implementation of the function $GetBestSamples$. Between different experiments, we only change the representation of sound supplied to the algorithm.

Implementation we use to obtain our results utilizes librosa [24] for the extraction of the audio features: Mel-spectrogram and MFCC. All features were extracted at a 22kHz sampling rate, with default parameters for the size of spectrogram frames (window sizes of 1024 and hop lengths of 512, Hamming window). The neural network was implemented and trained in Pytorch [25]. Matrix and vector computations are performed in NumPy [26], and for more computationally expensive matrix operations (distance calculations for determining the nearest neighbor) we also use Pytorch. On a system with an NVIDIA 2080Ti GPU, the use of GPU for distance calculations results in a significant speedup of approximately 2x when processing the entire dataset for evaluation.

For the encoder neural network in both BYOL and autoencoder approaches, we use a 5-layer convolutional neural network with kernel size 3 and 128 channels in each layer. For the decoder module in autoencoder, and the projection and predictor modules of BYOL, we found 2 layer convolutional networks sufficient, and adding layers to those modules resulted in no improvements. The network is trained for 1000 iterations with Adam optimizer and 256 batch sizes.

For BYOL augmentations we have tested the following: addition of Gaussian noise (mean 0, standard devaitaion 0.3),randomly removing 20% of the 128 Mel-spectrogram frequency bins, transposition by a fixed number of mel frequency bins and an augmentation based on Harmonic-Percussive Sound Separation (HPSS). The last augmentation applies HPSS to separate the harmonic components from one of the samples in pair (x, x') in the loss function of BYOL (Eq. 1). Final results are obtained with a combination of all agumentations.

As our end goal is to find all occurrences of the sound, the key figure of merit is recall, and as the task is practically oriented, our evaluation is based on the recall achievable within given $patience$. Precision of the system is less relevant, as $patience$ directly limits how many false-positive answers can occur.

To contextualize the following results, we measured the results of a naive nearest neighboor approach without the use of active learning. The naive baseline achieves a recall of 0.3 on the Samples subset of the data, and 0.51 on the Loops subset of the data.

A. Results In The Samples Category

Results in the audio samples category are shown in Fig. 1, including recall over specific genres. A learned feature extractor outperforms the standard MFCC feature extractor. The improvements are seen mainly in Drum&Bass and Downtempo songs. Within the House genre, retrieval achieves equally good results for both approaches, which can be largely explained by the low structural complexity of songs in this part of the dataset (several tracks have a single repeating loop as a baseline for the entire track, which makes retrieval significantly easier). We can see that the main difficulty appears in the Dubstep genre, where feature learning achieves no improvement.

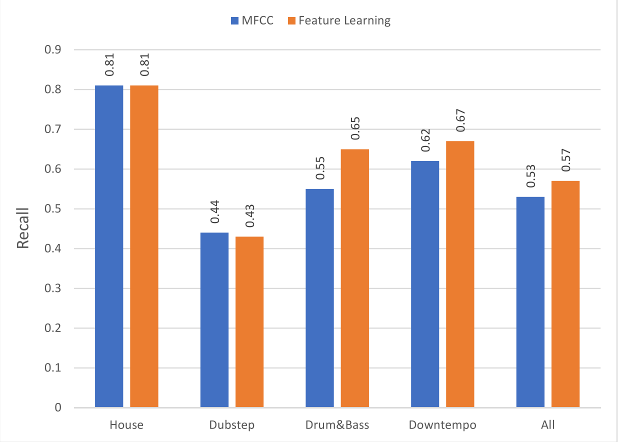


Fig. 1. Results in the *sound samples* category

B. Results In The Loops Category

Results in the audio loops category are shown in Fig. 2, including recall over specific genres. We can see that the retrieval of loops can be significantly easier than finding audio samples, likely stemming from the fact that by our definition loops share melodic and rhythmic qualities while for samples, the musical background can vary a lot. The feature learning approach improves over the MFCC features in overall results, and the improvements are seen within every genre. Much like in the sample searching task, the highest performance is seen in the House genre, and Dubstep is an outlier in being significantly harder than other genres.

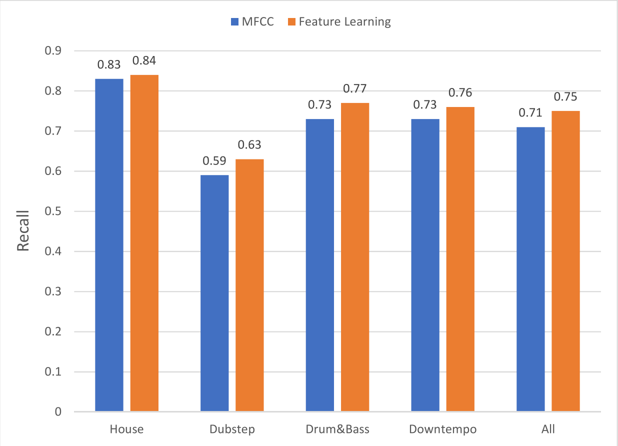


Fig. 2. Results in the *loops* category

C. BYOL vs. Autoencoder

Fig. 3 shows the comparison of results depending on the chosen loss function in both samples and loop categories, which motivated us to choose a combination of BYOL and autoencoder loss. While the BYOL loss alone is insufficient and autoencoder is enough to outperform MFCC features, the best performance is achieved when using a combination of both approaches.

D. Augmentation Choice

Fig. 4 shows the comparison of results depending on the choice of augmentation in BYOL training. The augmentation selection for BYOL ended up being less crucial than expected. This is especially seen for Harmonic-Percussive Separation, which we expected to improve the results by helping the extractor to focus on respectively percussive (more significant for sound samples) or melodic (more significant for loops) components of the sound. In practice, simple augmentations such as dropping frequency bins or adding Gaussian noise are enough to improve the results over autoencoder alone and the use of highly computationally complex HPSS isn't justified by the results.

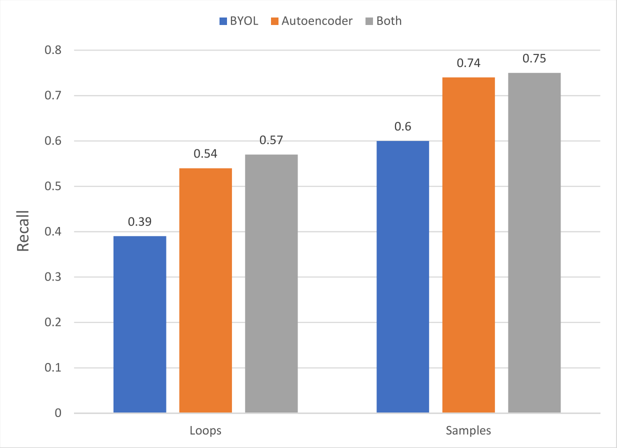


Fig. 3. Reuslts deppending on feature learning approach

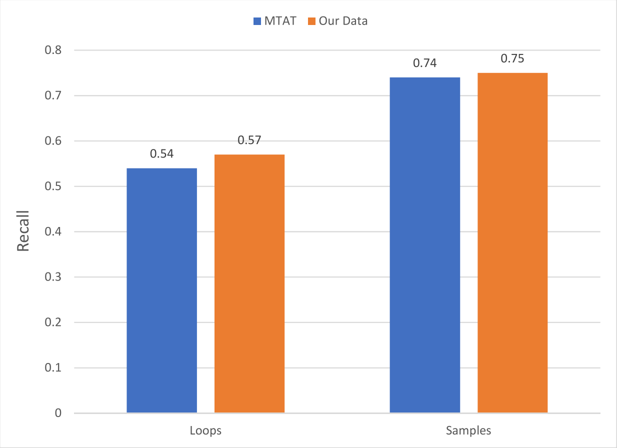


Fig. 5. Comparison of training the feature extractor on MTAT and our dataset

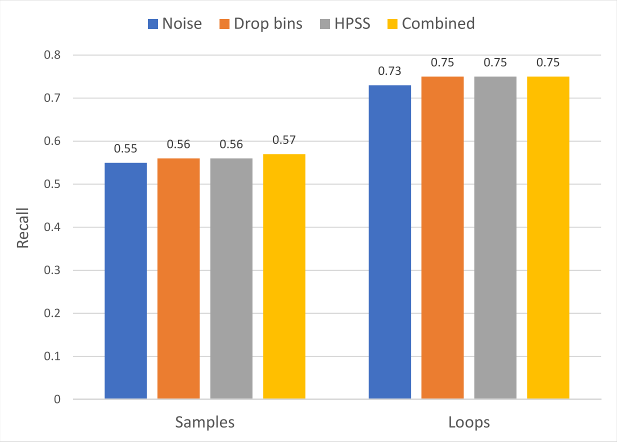


Fig. 4. Results depending on augmentations used in BYOL

E. More Representative vs. Bigger Training Data

As the key problem that motivates the use of feature learning is limited data availability, we also compare the results of training on a larger available dataset against the use of a small, but representative dataset. For this comparison, we used our sampleswap.org evaluation data (without labels) to train the "representative" model, while the "bigger" model is trained on a set of 15000 songs from the MagnaTagATune dataset. MagnaTagATune includes electronic music, but is not focused on it, and contains many genres irrelevant for our evaluation set such as classical and folk music. In Fig. 5, we compare the results. As can be seen, we achieve a similar retrieval quality with both approaches, but hand-selected representative data slightly outperforms training on a larger, but non-representative dataset.

V. Conclusions And Future Work

We have demonstrated a feature learning approach improving on our earlier work on the retrieval of subjectively interesting sound excerpts. The task concerns using active learning and user interaction to find multiple occurrences of an interesting sound in a music piece.

To facilitate a better evaluation of our results, we have developed our dataset to separate two categories of potential interesting "sound components" of electronic music. *Loops* are repeating melodic and rhythmic patterns commonly used by electronic music composers, while *sound samples* are sounds that stand out against the musical background and do not have to be melodic or rhythmic in nature. We have found that *sound samples* are significantly harder to find using our approach and may require further effort to separate well from the musical background.

The learneble feature extractor we used was a neural network trained in a fully unsupervised manner, using the principles of autoencoding and contrastive learning. We have found that this approach improves results when compared to a MFCC representation. A more detailed examination of the method's performance shows a number of conclusions. For best performance, autoencoding and BYOL approaches fto feature learning can be combined. Neither of these achieves the best results alone. In BYOL, an agumentation-based contrastive learning approach, the choice of augmentation affected the results, but the effect was not crucial to achieving good results. We have found that training on a small, but representative dataset was better than using a larger dataset with wide variety of music.

Future directions of development could include improvements of the neural network architecture, and development of the unsupervised learning method to achieve a representation that better separates ditinct sound components.

VI. Acknowledgments

This research was carried out in cooperation with Scalac sp. z o.o. and Vixa Games game development team as a part of the *Elaboration and implementation of audio modules (audio tracks indexing and analysis) and video modules (visualization of processed data) involving unique sound and* *graphic design use for the multimedia applications requirements* RPPM.01.01.01-22-0011/17 project co-funded by the European Union.

References

[1] E. J. Humphrey, J. P. Bello, Y. LeCun, “Moving beyond feature design: Deep architectures and automatic feature learning in music informatics,” in *ISMIR 2012*, pp. 403-408.

[2] M. Defferrard, K. Benzi, P. Vandergheynst, X. Bresson, “FMA: A dataset for music analysis,” arXiv preprint arXiv:1612.01840. 2017, https://doi.org/10.48550/arXiv.1612.01840

[3] Y. A. Chen, Y. H. Yang, J. C. Wang, H. Chen, “The AMG1608 dataset for music emotion recognition,” in *ICASSP 2015*, pp. 693-697, https://doi.org/0.1109/ICASSP.2015.7178058

[4] J. W. Kim, J. Salamon, P. Li, J. P. Bello, “Crepe: A convolutional representation for pitch estimation,” in *ICASSP 2018*, pp. 161-165, https://doi.org/10.1109/ICASSP.2018.8461329

[5] J. Jakubik, “Retrieving Sound Samples of Subjective Interest With User Interaction,” in *Proc. of the 2020 Federated Conference on Computer Science and Information Systems*, 2020, pp. 387-390, https://doi.org/10.15439/2020F82

[6] B. McFee, D. Ellis, “Analyzing Song Structure with Spectral Clustering,” in *ISMIR 2014*, pp. 405-410, https://doi.org/10.5281/zenodo.1415778

[7] Kothinti, S., Imoto, K., Chakrabarty, D., Sell, G., Watanabe, S., Elhilali, M. (2019, May). “Joint acoustic and class inference for weakly supervised sound event detection,” in *ICASSP 2019*, pp. 36-40, https://doi.org/10.1109/ICASSP.2019.8682772

[8] H. Xie, T. V. Huang, “Zero-Shot Audio Classification via Semantic Embeddings,” in *IEEE/ACM Transactions on Audio, Speech, and Language Processing,* vol. 29, 2021, pp. 1233-1242, https://doi.org/10.48550/arXiv.2011.12133

[9] S. Makino, "Audio source separation," Springer, 2018.

[10] J. P. Bello, L. Daudet, S. Abdallah, C. Duxbury, M. Davies, M. B. Sandler, “A tutorial on onset detection in music signals,” in *IEEE Transactions on speech and audio processing*, vol. 13, no. 5, 2005, pp. 1035-1047, https://doi.org/10.1109/TSA.2005.851998

[11] R. Marxer, J. Janer, "Study of Regularizations and Constraints in NMF-Based Drums Monaural Separation", in *Proc. of the 7th Int. Conference on Digital Audio Effects (DAFx’13).* Maynooth, Ireland, 2013.

[12] L. Lu, M. Wang, H. J. Zhang, “Repeating pattern discovery and structure analysis from acoustic music data,” in *Proc. of the 6th ACM SIGMM Int. Workshop on Multimedia Information Retrieval*, 2016, pp. 275-282, https://doi.org/10.1145/1026711.1026756

[13] P. López-Serrano, C. Dittmar, J. Driedger, M. Müller, “Towards Modeling and Decomposing Loop-Based Electronic Music,” in *ISMIR 2016*, pp. 502-508.

[14] J. B. L. Smith, M. Goto, “Nonnegative tensor factorization for source separation of loops in audio,” in *ICASSP 2018*, Calgary, Canada, pp. 171–175, https://doi.org/10.1109/MSP.2018.2877582

[15] J. B. L. Smith, Y. Kawasaki, M. Goto, “Unmixer: An interface for extracting and remixing loops,” in *ISMIR 2019*, Delft, Nethedlands, pp. 824–831, https://doi.org/10.5281/zenodo.3527938

[16] C. Chen, S. Xin, “Combined Transfer and Active Learning for High Accuracy Music Genre Classification Method,” in *2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE)*, IEEE, 2021, https://doi.org/10.1109/ICBAIE52039.2021.9390062

[17] A. Sarasúa, C. Laurier, P. Herrera, “Support vector machine active learning for music mood tagging,” in *9th International Symposium on Computer Music Modeling and Retrieval (CMMR)*, London, 2012, https://doi.org/10.1007/s00530-006-0032-2

[18] W. Li, X. Feng, M. Xue, “Reducing manual labeling in singing voice detection: An active learning approach,” in *2016 IEEE International Conference on Multimedia and Expo (ICME)* IEEE, 2016, https://doi.org/10.1109/ICME.2016.7552987

[19] Fu, Yifan, Xingquan Zhu, and Bin Li. “A survey on instance selection for active learning,” in *Knowledge and information systems*, vol. 35.2, pp. 249-283, 2013, https://doi.org/10.1007/s10115-012-0507-8

[20] T. H. Hsieh, L. Su, Y. H. Yang, “A streamlined encoder/decoder architecture for melody extraction,” in *ICASSP 2019*, pp. 156-160, https://doi.org/10.1109/ICASSP.2019.8682389

[21] J. Spijkervet, J. A.Y. Burgoyne, "Contrastive Learning of Musical Representations." arXiv preprint arXiv:2103.09410, 2021, https://doi.org/10.48550/arXiv.2103.09410

[22] Grill, J. B., Strub, F., Altché, F., Tallec, C., Richemond, P., Buchatskaya, E., Valko, M. (2020). Bootstrap your own latent-a new approach to self-supervised learning. Advances in Neural Information Processing Systems, 33, 21271-21284, https://doi.org/10.48550/arXiv.2006.07733

[23] Nguyen, K., Nguyen, Y., & Le, B. (2021). Semi-Supervising Learning, Transfer Learning, and Knowledge Distillation with SimCLR. arXiv preprint arXiv:2108.00587, https://doi.org/10.48550/arXiv.2108.00587

[24] B. McFee, C. Raffel, D. Liang, D. P. W. Ellis, M. McVicar, E. Battenberg, O. Nieto, “librosa: Audio and music signal analysis in python,” in *Proc. of the 14th python in science conference*, pp. 18-25, 2015.

[25] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, et al. “PyTorch: An Imperative Style, High-Performance Deep Learning Library,” in *Advances in Neural Information Processing Systems*, vol. 32, 2019, pp. 8024-8035, https://doi.org/10.48550/arXiv.1912.01703

[26] C.R. Harris, K.J. Millman, S.J. van der Walt, “Array programming with NumPy,” *Nature* vol. 585, pp. 357–362, 2020. DOI: 0.1038/s41586-020-2649-2, https://doi.org/10.1038/s41586-020-2649-2

[27] F. Pedregosa et al., “Scikit-learn: Machine Learning in Python,” in *Hournal of Machine Learning Research*, vol. 12, pp. 2825-2830, 2011, https://doi.org/10.48550/arXiv.1201.0490

Position Papers of the of the 17th Conference on Computer Science and Intelligence Systems pp. 19–26

DOI: 10.15439/2022F222
ISSN 2300-5963 ACSIS, Vol. 31

The Compositional Rule of Inference vs the Bandler-Kohout Subproduct: a Comparison of Two Standard Rules of Inference

Katarzyna Miś, Michał Baczyński
University of Silesia in Katowice
Bankowa 14, 40-007 Katowice, Poland
Email: {katarzyna.mis, michal.baczynski}@us.edu.pl

***Abstract*—This contribution focuses on the most popular scheme of reasoning in approximate reasoning, generalized modus ponens. Also, we consider the case when the reasoning is performed with one fuzzy rule. Usually, the compositional rule of inference introduced by Zadeh is involved. However, it is also common to use the Bandler-Kohout subproduct. We compare these two rules showing by experimental results the conditions when applying one of them is more appropriate. We concentrate on an example of image transformation where applying a different rule of inference gives a different conclusion. Moreover, we point out some theoretical justifications for particular fuzzy connectives used in both methods (fuzzy implication functions, triangular norms and, in general, fuzzy conjunctions).**

I. Introduction

WHENEVER we have imprecise data but would like to obtain meaningful results, we use methods called approximate reasoning. In this contribution, we analyse approximate reasoning based on fuzzy sets regarding one scheme, generalised modus ponens. For this scheme, we infer using the following idea,

RULE:	IF x is A, THEN y is B
FACT:	x is A'
CONCLUSION:	y is B'

where A, A', B, B' are fuzzy sets representing properties of objects x and y. A and A' are such that they are only slightly different (in some subjective opinions and using this informal language). It is why the conclusion expressed by a B' should also be "similar" to B to keep the intention of approximate reasoning. In our investigations, we consider two rules of inference:

- the Compositional Rule of Inference (CRI), see [1]

$$B'(y) := \sup_{x\in X} T(A'(x), I(A(x), B(y))), \quad y \in Y, \tag{CRI}$$

- the Bandler-Kohout Subproduct (BKS), see [2]

$$B'(y) := \inf_{x\in X} I(A'(x), T(A(x), B(y))), \quad y \in Y, \tag{BKS}$$

where T is a t-norm or a generalization of a conjunction and I is a fuzzy implication. We analyse particular sample data in order to show when (CRI) is better than (BKS) and vice versa. It should be noted that various scientists study these two rules of inference, see, e.g. [3], [4]. We focus on image processing and show that using a different inference rule gives a distinct conclusion, what is reflected in the output image. Our main hypothesis is: if A and A' are quite "similar", then B and B' will be more similar when B' is obtained from (CRI). However, if A and A' are "different", then B and B' will be more similar when B' will be calculated from (BKS).

The paper is organised as follows. Section 2 recalls some necessary definitions and facts used in the sequel. In Section 3, we present some experimental results and state our conclusions, observations, and verifications of hypothesises. Section 4 presents some theoretical results that partially justify our assumptions.

II. Preliminaries

First, let us introduce a symbol $\mathcal{F}(X)$ as a family of all fuzzy sets on X. Let us start with recalling some standard definitions and facts regarding t-norms and fuzzy implications.

Definition 2.1 (see [5], [6]): A function $T\colon [0,1]^2 \to [0,1]$ is called a triangular norm (t-norm in short), if it satisfies the following conditions for all $x, y, z \in [0,1]$

(T1) $T(x,y) = T(y,x)$,
(T2) $T(x, T(y,z)) = T(T(x,y), z)$,
(T3) $T(x,y) \le T(x,z)$ for $y \le z$, i.e., $T(x,\cdot)$ is non-decreasing,
(T4) $T(x,1) = x$.

Theorem 2.2 (see [6, Theorem 5.1]): For a function $T\colon [0,1]^2 \to [0,1]$ the following statements are equivalent:

(i) T is a continuous Archimedean t-norm.
(ii) T has a continuous additive generator, i.e., there exists a continuous, strictly decreasing function $f\colon [0,1] \to [0,\infty]$ with $f(1) = 0$ such that

$$T(x,y) = f^{-1}\left(\min\{f(x)+f(y), f(0)\}\right), \quad x,y \in [0,1].$$

Moreover, such a representation is unique up to a positive multiplicative constant.

We will need the following characterization of convex functions.

Theorem 2.3 (see [7, Theorems 7.3.2 and 7.3.3]): If a function $f\colon [0,1] \to \mathbb{R}$ is continuous, then the following statements are equivalent:

(i) f is convex.

(ii) For all $x, y \in [0,1]$ such that $y \le x$ and all $\varepsilon > 0$ such that $x+\varepsilon, y+\varepsilon \in [0,1]$ it holds

$$f(y+\varepsilon) - f(y) \le f(x+\varepsilon) - f(x). \tag{1}$$

In our investigations we also use fuzzy implication functions.

Definition 2.4 (see [5], [8]): A function $I\colon [0,1]^2 \to [0,1]$ is called a fuzzy implication, if it satisfies the following conditions:

(I1) I is non-increasing with respect to the first variable,
(I2) I is non-decreasing with respect to the second variable,
(I3) $I(0,0) = I(1,1) = 1$ and $I(1,0) = 0$.

Definition 2.5 (see [8]): We say that a fuzzy implication I satisfies

(i) the identity principle, if

$$I(x,x) = 1, \quad x \in [0,1], \tag{IP}$$

(ii) the left neutrality property, if

$$I(1,y) = y, \quad y \in [0,1], \tag{NP}$$

(iii) the ordering property, if

$$x \le y \iff I(x,y) = 1, \quad x, y \in [0,1]. \tag{OP}$$

Definition 2.6 (see [8, Definition 2.5.1]): A function $I\colon [0,1]^2 \to [0,1]$ is called an R-implication if there exists a t-norm T such that

$$I(x,y) = \sup\{t \in [0,1] \mid T(x,t) \le y\}, \qquad x, y \in [0,1]. \tag{2}$$

If I is generated from a t-norm T, then it will be denoted by I_T.

For R-implications generated from left continuous t-norms we have the following characterization.

Theorem 2.7 (cf. [8, Proposition 2.5.2]): For a t-norm T the following statements are equivalent:

(i) T is left-continuous.
(ii) A pair (T, I_T) satisfies the following residual principle

$$T(x,z) \le y \iff I_T(x,y) \ge z, \quad x, y, z \in [0,1], \tag{RP}$$

(iii) The supremum in the formula (2) is the maximum, i.e.,

$$I_T(x,y) = \max\{t \in [0,1] \mid T(x,t) \le y\}, \quad x, y \in [0,1]. \tag{3}$$

Theorem 2.8 (see [8, Theorem 2.5.21]): If T is a continuous Archimedean t-norm with the additive generator f as given in Theorem 2.2, then

$$I_T(x,y) = f^{-1}(\max\{f(y) - f(x), 0\}), \quad x, y \in [0,1]. \tag{4}$$

III. Experimental results

Here, as we mentioned in the Introduction, we will consider the case when our set of fuzzy rules contains only one rule. Therefore the inference process will proceed exactly according to (CRI) and (BKS). Let us take two different rules that concern the same topic. In both cases we will use (CRI) and (BKS) and we will compare our results.

In this matter, we would like to compare fuzzy sets $A \in \mathcal{F}(X)$ and $B \in \mathcal{F}(Y)$. It is important to show that dependencies between them have an influence on a choice of the rule of inference (CRI) or (BKS). Keeping in mind $X \neq Y$, we cannot calculate the standard similarity measure. However, we will use this notion to construct a function which compares A and B. Throughout literature we may find different properties of similarity measures and in a consequence different sets of axioms (see [9]–[12]). Let us mention some of them which can be considered here. Let $S\colon \mathcal{F}(X)^2 \to [0,1]$.

(P1) $S(X,\emptyset) = 0,\ S(A,A) = 1, \quad A \in \mathcal{F}(X)$,
(P2) $S(A,B) = S(B,A), \quad A, B \in \mathcal{F}(X)$,
(P3) $S(A,B) = S(A_\sigma, B_\sigma)$, $A, B \in \mathcal{F}(X)$, where if $A = [a_1, \dots, a_n]$, $B = [b_1, \dots, b_n]$ then $A_\sigma = [a_{\sigma(1)}, \dots, a_{\sigma(n)}]$, $B_\sigma = [b_{\sigma(1)}, \dots, b_{\sigma(n)}]$ and $\sigma \in S_n$ (is a permutation of {1, ..., n}).

Let us take the following two well-known similarity measures (see [9] and [12]),

$$M_1(A,B) = \begin{cases} 1, & A = B = \emptyset, \\ \frac{\sum_{i=1}^n \min\{A(a_i), B(b_i)\}}{\sum_{i=1}^n \max\{A(a_i), B(b_i)\}}, & \text{otherwise}, \end{cases}$$

and

$$M_2(A,B) = AM_{i=1}^n (1 - |a_i - b_i|),\ A, B \in \mathcal{F}(X),$$

where AM is the arithmetic mean.
For comparing sets $A \in \mathcal{F}(X)$, $B \in \mathcal{F}(Y)$ we will assume that $|X| = |Y| = n$ and take the following function $N_1, N^1, N_2, N^2\colon \mathcal{F}(X) \times \mathcal{F}(Y) \to [0,1]$.

$$N_1(A,B) = \begin{cases} 1, & A = B = \emptyset, \\ \min\limits_{\sigma,\tau \in S_n} \frac{\sum_{i=1}^n \min\{A(a_{\sigma(i)}), B(b_{\tau(i)})\}}{\sum_{i=1}^n \max\{A(a_{\sigma(i)}), B(b_{\tau(i)})\}}, & \text{otherwise}, \end{cases}$$

$$N^1(A,B) = \begin{cases} 1, & A = B = \emptyset, \\ \max\limits_{\sigma,\tau \in S_n} \frac{\sum_{i=1}^n \min\{A(a_{\sigma(i)}), B(b_{\tau(i)})\}}{\sum_{i=1}^n \max\{A(a_{\sigma(i)}), B(b_{\tau(i)})\}}, & \text{otherwise}, \end{cases}$$

$$N_2(A,B) = \min_{\sigma,\tau \in S_n} AM_{i=1}^n (1 - |a_{\sigma(i)} - b_{\tau(i)}|),$$

$$N^2(A,B) = \max_{\sigma,\tau \in S_n} AM_{i=1}^n (1 - |a_{\sigma(i)} - b_{\tau(i)}|),$$

where $A \in \mathcal{F}(X), B \in \mathcal{F}(Y)$. For these functions we see that for instance $G(X,\emptyset) = 0 = G(\emptyset, Y)$ and $G(A,A) = 1$, where $A = [x_1, \dots, x_n] = [y_1, \dots, y_n]$ and $G \in \{N_1, N^1, N_2, N^2\}$. Symmetry cannot be checked because of the domain (which is not symmetric, however if we take a function defined on a domain $\mathcal{F}(Y) \times \mathcal{F}(X)$, then of course values will be equal).

Now let us consider two examples which can show the guidelines for the choice between (CRI) and (BKS). First, let us mention that our motivation for this work is the comparison of two different images which were obtained as the conclusions from (CRI) (Fig. 2) and (BKS) (Fig. 3) - each pixel was considered as an input with one fuzzy rule used in the inference process (see [13]).

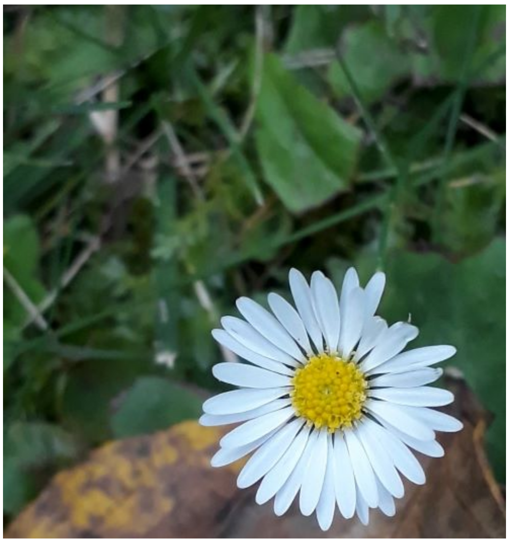


Fig. 1. The original image.

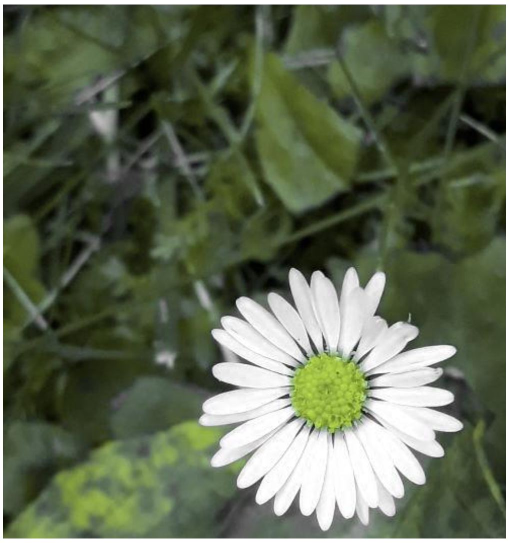


Fig. 2. Image obtained with (CRI).

In the following examples we used NumPy and Matplotlib libraries for Python (see [14], [15]). Also we have applied two pairs of (T, I):

1) $(T_{\mathbf{P}}, I_{\mathbf{GG}})$, where $T_{\mathbf{P}}(x,y) = xy$ and

$$I_{\mathbf{GG}}(x,y) = I_{T_{\mathbf{P}}}(x,y) = \begin{cases} 1, & x \leq y, \\ \frac{y}{x}, & x > y, \end{cases}$$

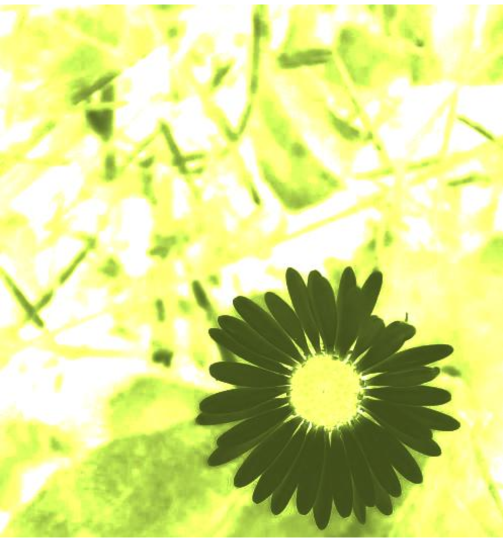


Fig. 3. Image obtained with (BKS).

2) $(T_{\mathbf{LK}}, I_{\mathbf{LK}})$, where $T_{\mathbf{LK}}(x,y) = \max\{0, x+y-1\}$ and $I_{\mathbf{LK}}(x,y) = I_{T_{\mathbf{LK}}}(x,y) = \min\{1, 1-x+y\}$.

In both these cases, we have left-continuous t-norms and R-implications generated by corresponding t-norms.

Example 3.1: This example is directly connected with the transformations of Fig. 1 which are presented above. However, because of the quite big size of the original image, we analyse the one consisting of small parts of it (Fig. 4). It contains different colours visible in the Fig. 1 and it has 1456 pixels.

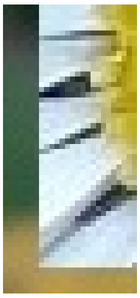


Fig. 4. Image made of pieces of the Fig. 1.

The first rule which is used by us is the following:

If an input pixel is 🟨 then an output pixel is 🟩

It means: if the pixel has values $[246, 246, 81]$, then the output pixel has values $[206, 249, 88]$. Then for fuzzy sets A, B representing these values of pixels (which in general can be from the different universes) we have $N^1(A,B) = 0.915, N_1(A,B) = 0.506, N^2(A,B) = 0.935, N_2(A,B) = 0.522$, so in all cases similarity is rather high. Now let us see how the similarity of A and A' looks like compared with the one of B and B'. The results are given in the following charts. To make the plots more clear we have drawn them for every second pixel from the Fig. 4.

We can see that regardless what similarity measure is used, the similarity of $B\&B'$ is directly proportional to the similarity of $A\&A'$ for the rule (CRI) (Fig. 5, 6, 9, 10). However in the case of (BKS) the situation is not as clear as before (see Fig. 7, 8, 11, 12). Nevertheless, we might say that for many input data the similarity of $B\&B'$ is inversely proportional, in particular to data where similarity of $A\&A'$ is greater than 0.5. These conclusions can be also confirmed by the linear regression (in magenta).

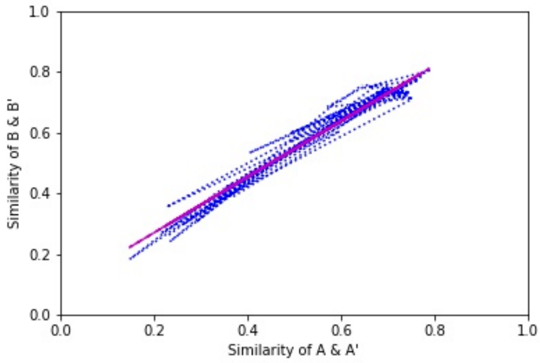


Fig. 5. Dependence between similarities calculated with M_1 for (CRI) and $(T_{\mathbf{P}}, I_{\mathbf{GG}})$, 1st rule.

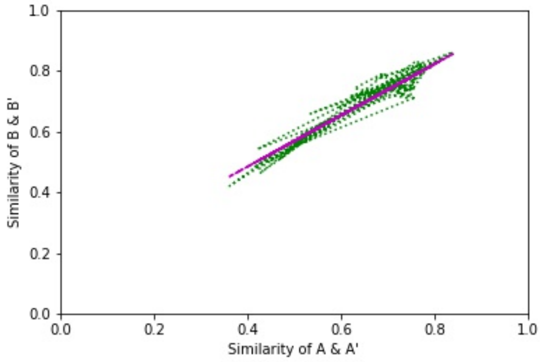


Fig. 6. Dependence between similarities calculated with M_2(CRI) and $(T_{\mathbf{P}}, I_{\mathbf{GG}})$, 1st rule.

Example 3.2: Here we consider the same Fig. 4, the same pairs $(T_{\mathbf{P}}, I_{\mathbf{GG}}), (T_{\mathbf{LK}}, I_{\mathbf{LK}})$ but we have another rule (we call it the 2nd rule):

If an input pixel is 🟨 then an output pixel is 🟪

It means: if the pixel has values $[246, 246, 81]$, then the output pixel has values $[128, 42, 239]$. Hence, for fuzzy sets A, B we have $N^1(A, B) = 0.714, N_1(A, B) = 0.346, N^2(A, B) = 0.785, N_2(A, B) = 0.372$, so in all cases similarity is lower than in Example 3.1. Now let us compare obtained similarities as we did before. On Figures 13, 14, 17, 18 we can see that values of similarities are still directly proportional. Simultaneously, we might say that for most of data obtained from BKS the similarity of $B\&B'$ is inversely proportional to the

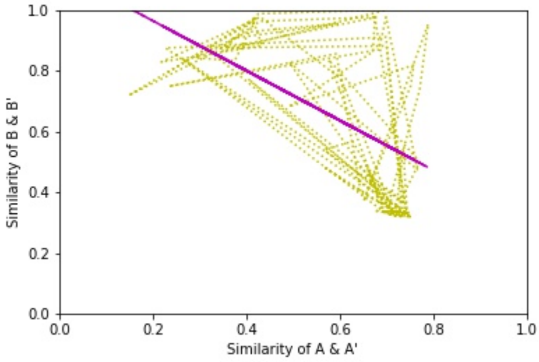


Fig. 7. Dependence between similarities calculated with M_1 (BKS) and $(T_{\mathbf{P}}, I_{\mathbf{GG}})$, 1st rule.

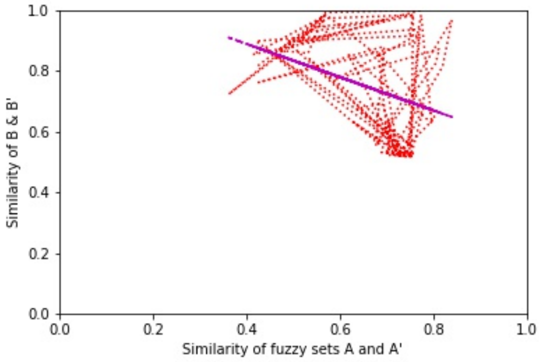


Fig. 8. Dependence between similarities calculated with M_2 (BKS) and $(T_{\mathbf{P}}, I_{\mathbf{GG}})$, 1st rule.

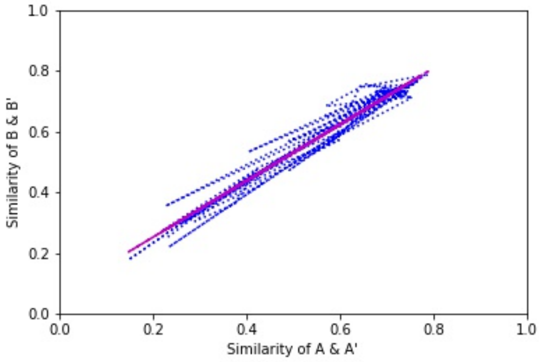


Fig. 9. Dependence between similarities calculated with M_1 for (CRI) and $(T_{\mathbf{LK}}, I_{\mathbf{LK}})$, 1st rule.

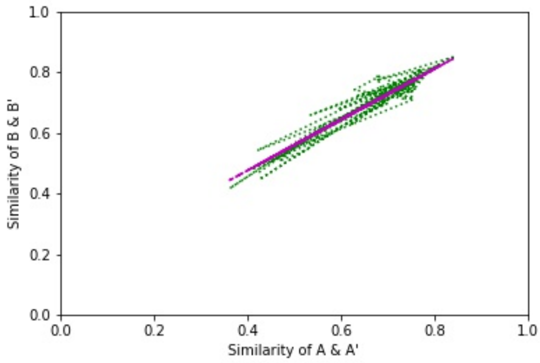


Fig. 10. Dependence between similarities calculated with M_2(CRI) and $(T_{\mathbf{LK}}, I_{\mathbf{LK}})$, 1st rule.

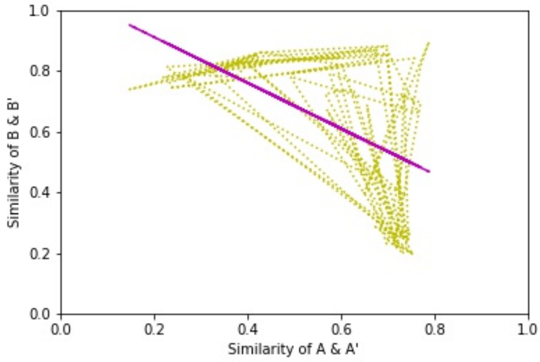


Fig. 11. Dependence between similarities calculated with M_1 (BKS) and $(T_{\mathbf{LK}}, I_{\mathbf{LK}})$, 1st rule.

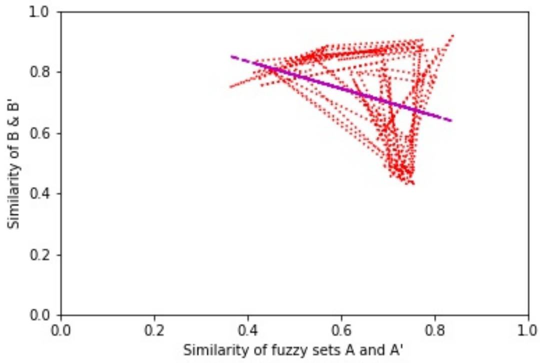


Fig. 12. Dependence between similarities calculated with M_2 (BKS) and $(T_{\mathbf{LK}}, I_{\mathbf{LK}})$, 1st rule.

similarities of $A\&A'$ (Fig. 15, 16, 19). Here the exception is only Figure 20, where we cannot say that.

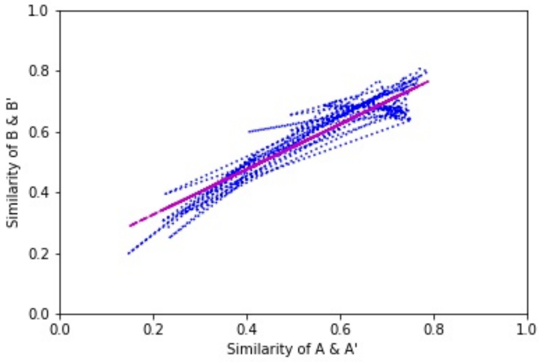


Fig. 13. Dependence between similarities calculated with M_1 for (CRI) and $(T_{\mathbf{P}}, I_{\mathbf{GG}})$, 2nd rule.

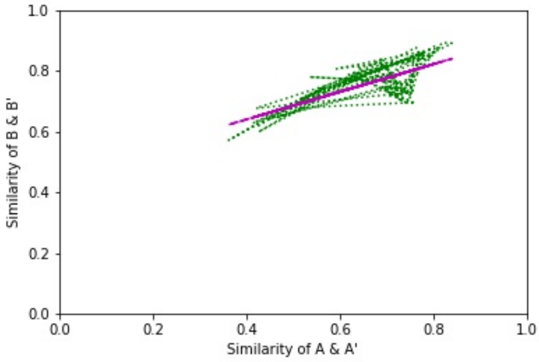


Fig. 14. Dependence between similarities calculated with M_2(CRI) and $(T_{\mathbf{P}}, I_{\mathbf{GG}})$, 2nd rule.

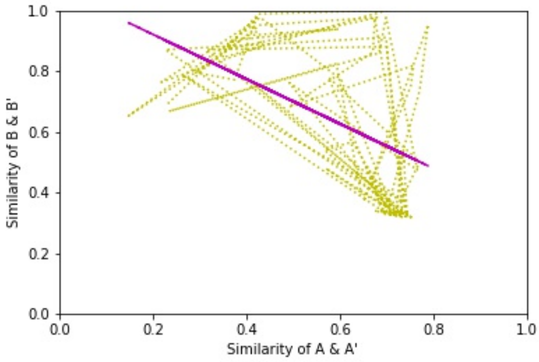


Fig. 15. Dependence between similarities calculated with M_1 (BKS) and $(T_{\mathbf{P}}, I_{\mathbf{GG}})$, 2nd rule.

After these examples we state the following observations, which are not what we expected at the beginning.

Hypothesis 1: The more A and B are similar, the more B and B' are similar for (CRI).

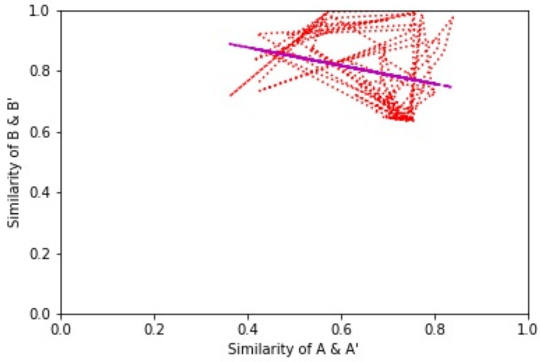


Fig. 16. Dependence between similarities calculated with M_2 (BKS) and $(T_{\mathbf{P}}, I_{\mathbf{GG}})$, 2nd rule.

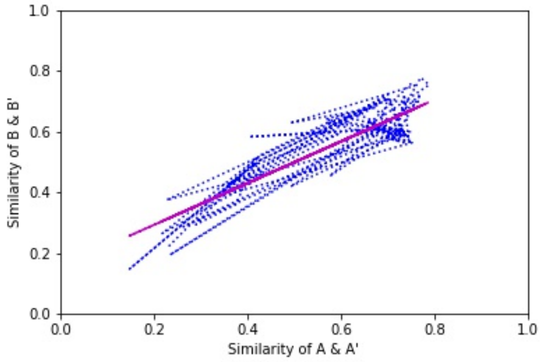


Fig. 17. Dependence between similarities calculated with M_1 for (CRI) and $(T_{\mathbf{LK}}, I_{\mathbf{LK}})$, 2nd rule.

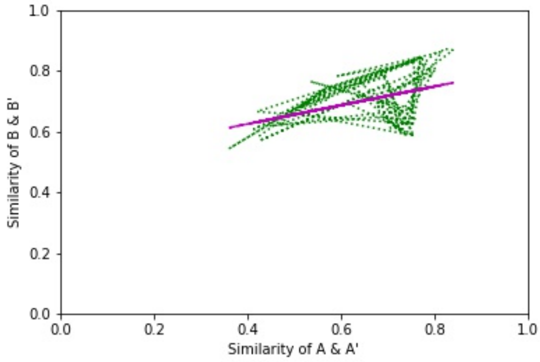


Fig. 18. Dependence between similarities calculated with M_2(CRI) and $(T_{\mathbf{LK}}, I_{\mathbf{LK}})$, 2nd rule.

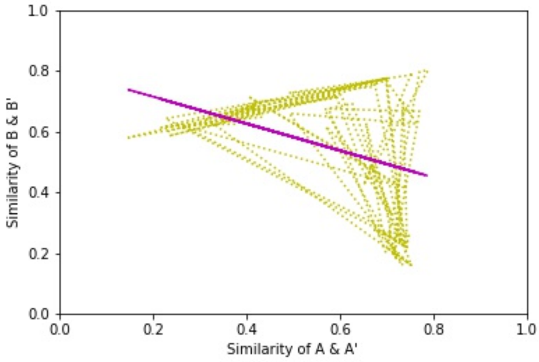


Fig. 19. Dependence between similarities calculated with M_1 (BKS) and $(T_{\mathbf{LK}}, I_{\mathbf{LK}})$, 2nd rule.

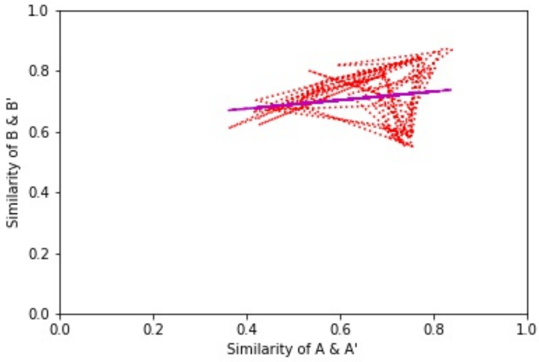


Fig. 20. Dependence between similarities calculated with M_2 (BKS) and $(T_{\mathbf{LK}}, I_{\mathbf{LK}})$, 2nd rule.

Hypothesis 2: The less A and B are similar, the more B and B' are similar for (BKS).

It turned out, it is not entirely true. Hence, our observation and conclusion are as follows.

Observation 1: Let $A, A' \in \mathcal{F}(X), B, B' \in \mathcal{F}(Y)$.

(i) The similarity of B and B' is directly proportional to the similarity of A and A' for the rule (CRI).

(ii) The similarity of B and B' is not always proportional to the similarity of A and A' for the rule (BKS).

(iii) The similarity of B and B' is usually inversely proportional to the similarity of A and A' for the rule (BKS).

IV. Theoretical Part

In this section, we want to justify the point (i) from Observation 1.

Let us consider the case of R-implications generated from left-continuous t-norms. First of all let us recall that such pairs (T, I_T), where T is a left-continuous t-norm, satisfy

$$y = \sup_{x \in [0,1]} T(x, I(x,y)), \quad y \in [0,1], \qquad \text{(CRI-GMP)}$$

which can be seen as a generalization of the property of the interpolativity (see [13]).

Let us focus on the formula (CRI). Our initial assumption is A and A' express the fact there is small difference between some property of an object x.
Let us suppose that $|X| = |Y| = n, n \in \mathbb{N}, n > 1$ and let us denote $A = [x_1, \dots, x_n], A' = [x'_1, \dots, x'_n], B = [y_1, \dots, y_n]$ and let

$$\begin{aligned} \varepsilon_i &= |x_i - y_i|, \quad i = 1, \dots, n, \\ \delta_i &= |x_i - x'_i|, \quad i = 1, \dots, n. \end{aligned}$$

Also suppose that if A and B are 'similar', then $|x_i - y_j| \geq \varepsilon_i,\ i \neq j$.
We will show that the following inequality holds for any Archimedean continuous t-norm T with a convex generator f (t-norms used for the experiments have convex generators),

$$y_i + c \leq T(x'_i, I_T(x_i, y_i)) \leq y_i + a,$$

for $a \in \{-\varepsilon_i - \delta_i, 0, -\varepsilon_i + \delta_i\}, c \in \{-\varepsilon_i - \delta_i, -\delta_i, -\varepsilon_i + \delta_i\}$ and x_i such that $x_i - \varepsilon_i - \delta_i \geq 0,\ i = 1, \dots, n$.
Firstly, let us consider the case $\delta_i = x_i - x'_i$.

1) if $x_i \leq y_i$, then we have

$$\begin{aligned} T(x_i - \delta_i, I_T(x_i, y_i)) &= T(x_i - \delta_i, 1) \\ &= x_i - \delta_i \\ &= y_i - \varepsilon_i - \delta_i, \end{aligned}$$

and

$$y_i - \varepsilon_i - \delta_i = x_i - \delta_i \leq T(x_i - \delta_i, I_T(x_i, y_i)).$$

2) if $x_i > y_i$, then

$$\begin{aligned} &T(x_i - \delta_i, I_T(x_i, y_i)) \leq y_i \iff \\ &\qquad I_T(x_i - \delta_i, I_T(x_i, y_i)) \geq I_T(x_i, y_i), \end{aligned}$$

which is true from the (RP) and the monotonicity of I_T. Now we will show $y_i - \delta_i \leq T(x_i - \delta_i, I_T(x_i, y_i))$. Let us recall inequality (1), which can rewritten in the following way

$$f(y + \varepsilon) + f(x) \leq f(x + \varepsilon) + f(y), \quad \text{where } y \leq x.$$

This can be applied here as

$$f(x_i - \delta_i) + f(x_i - \varepsilon_i) \leq f(x_i) + f(x_i - \varepsilon_i - \delta_i),$$

where $x := x_i - \varepsilon_i, \varepsilon = \varepsilon_i, y := x_i - \varepsilon_i - \delta_i$. The above inequality is equivalent to

$$\begin{gathered} f(x_i - \delta_i) + f(x_i - \varepsilon_i) - f(x_i) \leq f(x_i - \varepsilon_i - \delta_i) \\ \iff \\ f^{-1}(f(x_i - \delta_i) + f(x_i - \varepsilon_i) - f(x_i)) \geq x_i - \varepsilon_i - \delta_i \end{gathered}$$

Note that $x_i - \varepsilon_i - \delta_i \geq 0$, so

$$f(x_i - \delta_i) + f(x_i - \varepsilon_i) - f(x_i) \leq f(0)$$

and

$$\begin{aligned} &\min\{f(0), f(x_i - \delta_i) + f(x_i - \varepsilon_i) - f(x_i)\} \\ &\qquad = f(x_i - \delta_i) + f(x_i - \varepsilon_i) - f(x_i), \end{aligned}$$

so further we may write

$$\begin{gathered} T(x_i - \delta_i, f^{-1}(f(x_i - \varepsilon_i) - f(x_i))) \geq x_i - \varepsilon_i - \delta_i \\ \iff \\ T(x_i - \delta_i, I_T(x_i, x_i - \varepsilon_i)) \geq x_i - \varepsilon_i - \delta_i \end{gathered}$$

Now, let $x'_i > x_i$, so $\delta_i = x'_i - x_i$.

1) if $x_i \leq y_i$, then we have

$$T(x'_i, I_T(x_i, y_i)) = x'_i = y_i + \delta_i - \varepsilon_i.$$

2) if $x_i > y_i$, then we have

$$\begin{aligned} T(x'_i, I_T(x_i, y_i)) &= T(x'_i, I_T(x'_i - \delta_i, x'_i - \delta_i - \varepsilon_i)) \\ &\leq x'_i, \end{aligned}$$

which, by (RP), is equivalent to

$$1 = I_T(x'_i, x'_i) \geq I_T(x_i - \delta_i, x'_i - \delta - \varepsilon_i)$$

and

$$T(x'_i, I_T(x_i, y_i)) \leq x'_i = y_i - \varepsilon_i + \delta_i.$$

Moreover,

$$x_i + \delta_i - \varepsilon_i \leq T(x'_i, I_T(x_i, y_i)).$$

Indeed, again using the property of convex continuous function from (1) we have

$$f(x + \varepsilon) + f(y) \geq f(x) + f(y + \varepsilon),$$

and applying it for the generator f of a t-norm T we obtain

$$f(x_i + 2\delta_i) + f(x_i - \varepsilon_i - \delta_i) \geq f(x_i + \delta_i) + f(x_i - \varepsilon_i),$$

for such substitutions:
$x := x_i + \delta_i$,
$y := x_i - \varepsilon_i - \delta_i$,
$\varepsilon := \delta_i$.
Next, from the fact f is strictly decreasing we may write $f(x_i) + f(x_i - \varepsilon_i - \delta_i) \geq f(x_i + 2\delta_i) + f(x_i - \varepsilon_i - \delta_i)$.
Therefore we have

$$f(x_i - \delta_i - \varepsilon_i) + f(x_i) \geq f(x_i + \delta_i) + f(x_i - \varepsilon_i),$$

that is equivalent to

$$\begin{gathered} f(x_i - \delta_i - \varepsilon_i) \geq f(x_i + \delta_i) + f(x_i - \varepsilon_i) - f(x_i) \\ \iff \\ x_i - \delta_i - \varepsilon_i \leq f^{-1}(f(x_i + \delta_i) + f(x_i - \varepsilon_i) - f(x_i)) \\ \iff \\ x_i - \delta_i - \varepsilon_i \leq T(x_i + \delta_i, f^{-1}(f(x_i - \varepsilon_i) - f(x_i))) \\ \iff \\ x_i - \delta_i - \varepsilon_i \leq T(x_i + \delta_i, I_T(x_i, x_i - \varepsilon_i)) \\ \iff \\ y_i - \delta_i \leq T(x'_i, I_T(x_i, y_i)) \end{gathered}$$

Here again, we used the fact that

$$f(x_i + \delta_i) + f(x_i - \varepsilon_i) - f(x_i) \leq f(x_i - \delta_i - \varepsilon_i) \leq f(0).$$

The conclusion is the following: for inferred $B' = [y'_1, \ldots, y'_n]$ values of y'_i for $i \in \{1, \ldots, n\}$ are in the neighbourhood of y_i and if ε_i, δ_i approach 0, y'_i also approaches y_i and in the consequence value of the similarity measure of B and B' is close to 1.

V. Conclusions

In this contribution, we have compared two rules of inference, the Compositional Rule of Inference and the Bandler-Kohout Subproduct. Our goal was to investigate some dependencies between input and output. Our observations are the following. The similarity of B and B' is directly proportional to the similarity of A and A' for the rule (CRI). The similarity of B and B' is not always proportional to the similarity of A and A' for the rule (BKS). The similarity of B and B' is usually inversely proportional to the similarity of A and A' for the rule (BKS), especially if the similarity of A and A' (the antecedent and the input) is greater than $0,5$. In future work, we want to study these methods deeply with more rules and for different fuzzy logical operations classes.

References

[1] L. A. Zadeh, "Outline of a new approach to the analysis of complex systems and decision processes," *IEEE Trans. Syst., Man, Cybern.*, vol. 3, pp. 28–44, 1973. doi: https://doi.org/10.1016/S0019-9958(65)90241-X

[2] W. Bandler and L. J. Kohout, *Fuzzy Relational Products as a Tool for Analysis and Synthesis of the Behaviour of Complex Natural and Artificial Systems*. Boston, MA: Springer US, 1980, pp. 341–367. ISBN 978-1-4684-3848-2

[3] M. Štěpnička and B. Jayaram, "On the Suitability of the Bandler-Kohout Subproduct as an Inference Mechanism," *IEEE Trans. Fuzzy Syst*, vol. 18, no. 2, pp. 285–298, 2010. doi: https://doi.org/10.1109/TFUZZ.2010.2041007

[4] S. Mandal and B. Jayaram, "Bandler-Kohout Subproduct with Yager's classes of Fuzzy Implications," *IEEE Trans. Fuzzy Syst.*, vol. 22, no. 3, pp. 469–482, 2014. doi: https://doi.org/10.1109/TFUZZ.2013.2260551

[5] J. Fodor and M. Roubens, *Fuzzy Preference Modelling and Multicriteria Decision Support*. Dordrecht: Kluwer Academic Publishers, 1994.

[6] E. P. Klement, R. Mesiar, and E. Pap, *Triangular Norms*. Dordrecht: Kluwer Academic Publishers, 2000. [Online]. Available: https://doi.org/10.1007/978-94-015-9540-7

[7] M. Kuczma, *An Introduction to the Theory of Functional Equations and Inequalities. Cauchy's Equation and Jensen's Inequality*. Warszawa, Kraków, Katowice: Państwowe Wydawnictwo Naukowe (Polish Scientific Publishers) and Uniwersytet Śląski, 1985.

[8] M. Baczyński and B. Jayaram, *Fuzzy Implications*, ser. Studies in Fuzziness and Soft Computing. Berlin Heidelberg: Springer, 2008, vol. 231.

[9] C. Pappis and N. Karacapilidis, "A comparative assessment of measures of similarity of fuzzy values," *Fuzzy Sets and Systems*, vol. 56, pp. 171–174, 1993. doi: https://doi.org/10.1016/0165-0114(93)90141-4

[10] J. Fan and W. Xie, "Some notes on similarity measure and proximity measure," *Fuzzy Sets and Systems*, vol. 101, pp. 403–412, 1999. doi: https://doi.org/10.1016/S0165-0114(97)00108-5

[11] I. Jenhani, S. Benferhat, and Z. Elouedi, *Possibilistic Similarity Measures*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 99–123. ISBN 978-3-642-10728-3

[12] Y. Li, K. Qin, and X. He, "Some new approaches to constructing similarity measures," *Fuzzy Sets and Systems*, vol. 234, pp. 46–60, 2014. doi: https://doi.org/10.1016/j.fss.2013.03.008

[13] K. Miś and M. Baczyński, "Some Remarks on Approximate Reasoning and Bandler-Kohout Subproduct," in *Information Processing and Management of Uncertainty in Knowledge-Based Systems*, ser. Communications in Computer and Information Science, M.-J. Lesot, S. Vieira, M. Reformat, J. Carvalho, B. Bouchon-Meunier, and R. Yager, Eds., vol. 1238. Springer, 2020. doi: https://doi.org/10.1007/978-3-030-50143-3_60 pp. 775–787.

[14] J. D. Hunter, "Matplotlib: A 2d graphics environment," *Computing in Science & Engineering*, vol. 9, no. 3, pp. 90–95, 2007. doi: https://doi.org/10.1109/MCSE.2007.55

[15] C. Harris, K. Millman, S. van der Walt, and et al., "Array programming with NumPy," *Nature*, vol. 585, pp. 357–362, 2020. doi: https://doi.org/10.1038/s41586-020-2649-2

DOI: 10.15439/2022F125
ISSN 2300-5963 ACSIS, Vol. 31

Prototypical investigation of the use of fuzzy measurement data in a case study in water analysis

Stefanie Penzel, Mathias Rudolph
Leipzig University of Applied Science
Karl-Liebknecht-Str. 132
04277 Leipzig, Germany
Email: {stefanie.penzel, mathias.rudolph}@htwk-leipzig.de

Helko Borsdorf
Helmholtz Centre for Environmental Research - UFZ,
Department of Monitoring and Exploration Technologies,
Permoserstr. 15 Leipzig,
Germany

Olfa Kanoun
University of Technology Chemnitz, Professorship Measurement and Sensor Technology, Reichenhainer Str. 9
Chemnitz, Germany

***Abstract*—A common problem when using real data is the fact that the values usually exhibit some degree of uncertainty. Measurement uncertainties therefore represent a major challenge when trying to interpret and draw conclusions from real data. This is especially true in on-site analysis in the environmental sector where the uncertainty in sample plays such a large role. An approach for the modelling and analyze of data for polluted water and the inclusion of measurement uncertainties is presented. This approach is based on fuzzy modelling, in which the uncertainty of the parameters is represented by so-called fuzzy numbers and thus reflect a possible blurred range of these parameter values. The result is a fuzzy pattern classifier, which allows a fuzzy and thus realistic characterization of unknown water samples. The procedure is exemplified using the extinction spectra taken using a UV/Vis spectrometer.**

I. Introduction

THE conservation of water resources and the need to continuously monitor the quality of these water resources (e.g., in watercourses, wastewater, bathing lakes, etc.) is of increasing importance nowadays. The determination of sufficient characteristic values to describe the water quality and the subsequent characterization represent a significant challenge. Various parameters play an important role in this. Polycyclic aromatic hydrocarbons (PAHs) such as benzene or naphthalene are a priority substance in water policy. In addition, other significant indicators may be relevant for the determination of pollutants. [1] These substances can be determined using a variety of standardized analytical methods. However, many methods have limitations, particularly when investigating very low concentrations in water. In addition, these measurement methods are traditionally performed in the laboratory mostly after water sampling at different locations at different times. These approaches are no longer considered efficient [2–4]. To detect and analyze the formation of pollutants directly at the source, an on-site sensor system is required. Continuous and unbiased measurements of this type can then be used for the optimal control and verification of water quality. For this reason, the Helmholtz Centre for Environmental Research (UFZ) is working on an experimental setup that can analyze water samples directly on-site using ultraviolet (UV) / visible (Vis) spectroscopy. Each sensor-based measurement has an objective uncertainty, which essentially depends on the measuring method and instrument. For example, for measuring instruments this uncertainty can be specified by an accuracy or error class according to DIN 1319-2, DIN 1319-3. Unstable operating conditions which occur especially in the environmental sector, and here with mobile on-site analysis, lead to additional uncertainties. For instance, seasonal changes in temperature and humidity may also contribute to the uncertainty of a measurement.

One of the advantages of using fuzzy classification methods is that such uncertainties can be characterized. The assignment to a pollutant substance is not crisply defined but is categorized according to a grade of membership. These are in the range between zero and one. The underlying fuzzy is based on a theory published by Zadeh in 1965 [5] and since then, it has been used and further developed in many areas, current such as the selection process for outsourcing users [6] or for the description of transportation problems through the extension of fuzzy sets [7]. The basic idea is to extend the classic binary classification, in this case the pollutant is present or not, to allow a gradual change. In our case, this allows the model to output that the pollutant may be present and further analysis is necessary. This is shown by the membership function to a fuzzy set. Such fuzzy forms of description, in which the crisp values are included as special cases, represent a new optimized meaning in the characterization of water quality. They have the advantages of greater flexibility and proximity to reality compared to the crisp forms of description, and moreover allow the adequate implementation of expert knowledge.

The procedure is demonstrated using the measurement data of water samples recorded by a UV/Vis test setup. The characteristic properties of different water samples (here measured in the form of extinction spectra) with different substance concentrations are to be derived from the data in a so-called learning phase and to be used to model a fuzzy pattern classifier. With this classifier, a statement is to be made as to whether a certain substance is present with uncertainty, present or present with a certain degree of certainty. If a substance is not present, the membership of these three classes will be zero. This serves as the basis for a fuzzy and, thus, realistic characterization of current (unknown) water samples in the next working phase. Depending on the results, recommendations for action can be made afterwards.

II. Methodical Basics

A. The UV/Vis measuring setup

The measuring device used is a UV/Vis measuring setup. UV/Vis spectroscopy uses electromagnetic radiation to detect substances in water. In the case of pollutants in water, for example, part of the radiation is absorbed by the pollutants. This can be seen in the absorption/extinction spectrum by comparing the measurement to a blank measurement. By means of the absorbed wavelength and the level of absorption, the type and concentration level of the substance can be concluded. The experimental setup was realized at the UFZ Leipzig and will be integrated into a mobile submersible probe for future on-site data acquisition. The setup consists of a UV/Vis light source with a deuterium and a tungsten lamp, a measuring cell, and a spectrometer. The connection of the three components is used via optical fibers. The measuring cell consists of two collimator adapters with optical windows, a stainless-steel flow cell and two 90° collimators each. The cuvette containing the dissolved sample substance can be inserted into the flow cell. The control and data acquisition of the measurement setup are carried out on a laptop using Python software, which automatically compares the measured data with a prerecorded blank spectrum to create an extinction spectrum from the two transmission spectra according to the Beer-Lambert law. Measurement uncertainties already arise during the testing in the laboratory, e. g. from the lamp due to fluctuations in energy supply or due to the noise caused by the spectrometer. For later mobile use, the results can also be affected by e. g. temperature or humidity. Which would be reflected in the noise behavior of the spectra, or it can lead to a rise in the baseline.

B. Fundamentals of the Fuzzy Pattern Classification

The fuzzy system used here is based on the fuzzy pattern classifier introduced by Bocklisch [8]. This methodology is widely used in pattern recognition for object classification. Here, a set of fuzzy membership functions μ: x ⟶ [0, 1] are created per class, which model characteristic features of this class. Through the membership functions, the feature values x ∈ IR of an object can be mapped to the unit interval, which represents the membership to a feature of an ideal class member. All the memberships are then merged and classified into the appropriate classes. Then the object is assigned to the class that has the highest aggregated value. This procedure is already used in many areas such as in signal processing applications and automation systems [9, 10] or in the field of neuronal statements and medical diagnostic reasoning [11, 12]. In addition, this fuzzy modeling is also used for data-inherent structures [13] or for online recognition of fuzzy time series patterns [14].

The exact procedure is divided into a ***learning phase*** and a ***working phase***. In the learning phase, a fuzzy classification model is constructed in a multidimensional feature space. This can be achieved by choosing between a data-driven or expert-based approach. In the data-driven procedure, several measurement runs are performed for predefined prototypical dilution series. First, the recorded object data sets (learning data) are divided into crisp groups. Two strategies are possible:

1) A cluster analysis (e.g., hierarchically agglomerative) is performed for the object data. This is a mathematical method, which creates corresponding groups through the accumulation of certain similar objects (in the sense of a small distance measure) as a result.

2) An a priori division of the objects into groups based on expert knowledge is carried out. This can be done by dividing the objects, here e.g., dividing the extinction maxima according to before or above the detection limit.

Subsequently, these crisp groups are transferred to fuzzy groups. The description of each group in the one- or multi-dimensional feature space is achieved here by a highly flexible, parametric membership function of the AIZERMAN potential function type. This function is described and illustrated in simplified form for the symmetrical one-dimensional case (Fig. 1).

$$\mu(u) = \frac{a}{1+\left(\frac{1}{b}-1\right)\cdot\left(\frac{|u-u_0|}{c}\right)^d} \qquad (1)$$

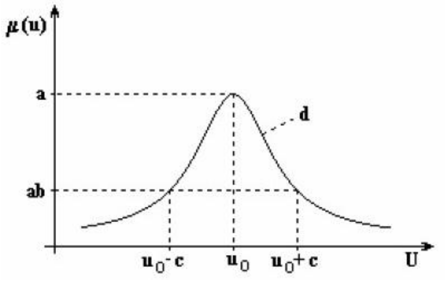


Fig. 1 One-dimensional AIZERMAN potential function [4]

Meaning of the parameters (see also Fig. 1):

- Local information u0 (crisp): representative of the fuzzy quantity (special case)

- Broadening c (to the left and right side of u0): precisely observed range
- Border membership b ϵ [0,1]: determines the membership values at the borders of the strictly area
- Maximum value of membership a (usually normalized with a = 1)
- d describes the continuously decreasing course of the membership function (d → ∞: crisp (binary) description)

The potential function can be used to describe both one-sided open intervals and closed intervals in a fuzzy way. In addition, the differentiation of the left and right-sided branch increases the adaptability. In the multivariate case, the characteristic dimension of the membership function is expanded accordingly, whereby each group can be represented by an analytically closed membership function. By using a closed analytical membership function to describe each group in the one- or multi-dimensional feature space, the method used here also differs from the rule-based fuzzy logic [15]. In a data-driven procedure, the parameters are calculated automatically from the recorded (learning) data sets by means of supervised learning [8].

The abovenamed AIZERMAN potential function may be applied to each axis of a multidimensional space. Thus, even information about high-dimensional groups can be described efficiently by a few parameters. A further advantage of the AIZERMAN potential function approach is that trapezoidal and triangular attribution functions as well as the so-called fuzzy singletons (crisp description as a special case of fuzzy case) can be converted into such a uniform description form, thus enabling a highly flexible and universal application with the possibility of modelling.

As an alternative to this data-driven approach, the parameters can be determined by expert knowledge, i. e for each of the characteristics fuzzy areas are defined manually and the fuzzy groups are then formed. This approach is typically used for linguisdetertic characteristics. Their values are not exactly defined, but colloquially defined by certain expressions (e. g. “small”, “medium”, “large”).

In the working phase, the classification model (Fuzzy Pattern Classifier) created in the learning phase is used for fuzzy identification of the current water sample (rep-resented by corresponding working data). The result is an membership or sympathy vector, whose components indicate the memberships to all declared classes. The current water status can be determined in a precise way from the maximum attribution values. The security (or risk) of this decision can be determined by the differences in the membership values.

III. Research results and discussion

A. *Structure of the data base*

Several dilution series with different concentrations of benzene, naphthalene, uranine and rhodamine B were prepared for the compilation of different data sets. By means of the measurement setup, extinction spectra were recorded for each substance at different concentrations (see Fig. 2).

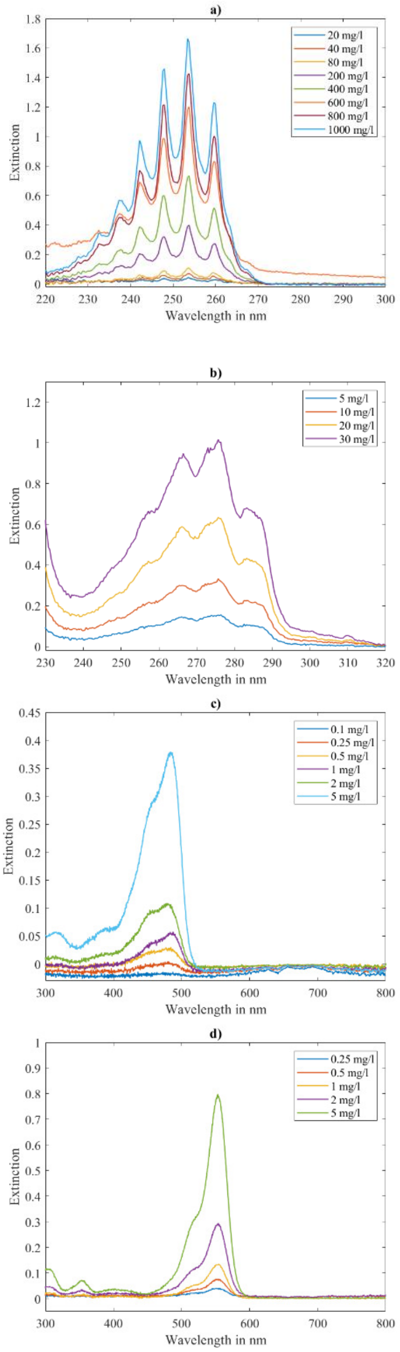


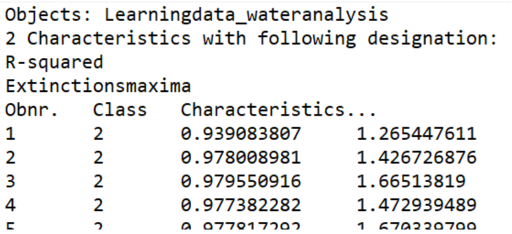
Fig. 2 Concentration-dependent extinction spectra of (a) benzene, (b) naphthalene, (c) uranine and (d) rhodamine B

Each substance was subjected to several measurement runs. Table 1 presents the selected dilution samples or their respective datasets and explains the abbreviations contained therein. The classification into this groups is based on the limit of detection or limit of quantification of the recorded data and is used for the subsequent classification into sharp classes.

TABLE I.
WATER SAMPLES USED WITH ABBREVIATED TITLE

Short description	Description of the samples with classification into the different existing classes
BUP	Benzene is unsurely present
BP	Benzene is present
BSP	Benzene is surely present
NUP	Naphthalene is unsurely present
NP	Naphthalene is present
NSP	Naphthalene is surely present
UUP	Uranine is unsurely present
UP	Uranine is present
USP	Uranine is surely present
RUP	Rhodamine B is unsurely present
RP	Rhodamine B is present
RSP	Rhodamine B is surely present

The extinction spectra of each compound were first described mathematically with an algorithm. For the mathematical description, several Gaussian functions were added to a total function and the parameters were each adapted to a spectrum of a substance. This characteristic overall function is then overlayed on all spectra and fitted to the spectra using the method of least squares. The R-squared is calculated. As the overall function is characteristic for each substance, the R-squared basically indicates the probability with which a certain substance is present. The extinction maximum gives a statement about the concentration content of the substance. These data were then stored in an overall dataset. Subsequently, the data sets were selected on a random basis and then divided into so-called learning and work data (see Section 2.1). The R-squared and extinction maxima were stored in an object file (.OTX). Fig. 3 shows a section of the created object file. This consists of a header with the necessary information about the data and then lists the object number, the corresponding class (here the assignment to the respective water sample) and the measured values for each of the two characteristics.



```
Objects: Learningdata_wateranalysis
2 Characteristics with following designation:
R-squared
Extinctionsmaxima
Obnr.   Class   Characteristics...
1       2       0.939083807     1.265447611
2       2       0.978008981     1.426726876
3       2       0.979550916     1.66513819
4       2       0.977382282     1.472939489
5       2       0.977817292     1.670339799
```

Fig. 3 Extract of learning data in OTX format

B. Classifier Development

In the ***learning phase***, a fuzzy classification model was first constructed in the two-dimensional characteristic space. The data-driven approach was combined with an expert-based approach by dividing the learning data into sharp groups and then building up the fuzzy pattern classifiers. A priori grouping is used to divide the learning data into crisp groups (see Fig. 4).

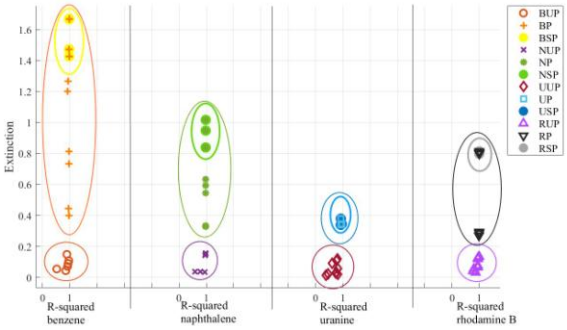


Fig. 4 Object distribution of data according to a priori grouping

The parameters for the fuzzy pattern classifier were transferred from crisp groups to fuzzy ones based on the parametric belonging function. For each characteristic of an object the membership function was described with the parameter values. Here, c is the elementary uncertainty of the objects and can be regarded as the measurement uncertainty of the respective measured values. Subsequently, the objects are first unified in one-dimensional sets and then transformed into multidimensional fuzzy pattern classes (in this case two-dimensional) using an N-fold compensatory Hamacher intersection operator [16]:

$$\cap_{Ham}^{N} \mu_i = \frac{1}{\frac{1}{n}\sum_{i=1}^{N}\frac{1}{\mu_i}} \qquad (2)$$

Here n describes the total numbers of dimensions and i present the index of the basis functions. If this is applied to all sets, the result is 12 classes in the two-dimensional feature space. (See Fig. 5).

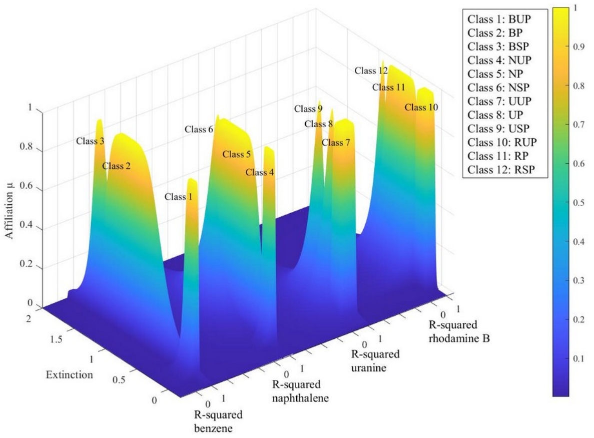


Fig. 5 Result of the classifier development

For this fuzzy pattern classifier, only the measured concentrations were considered. To ensure that very high concentrations can also be automatically included in the evaluation, the fuzzy pattern classifier was adapted using expert knowledge. The expert-based procedure offers a supplement to the data-driven approach, with which an adaptation of the constructed classifier can be carried out. Since no very high concentrations were measured in the present test runs, but it is known from Beer-Lambert's law that these also increase with increasing extinction maxima, the respective classes were enlarged in the direction of the feature "extinction maxima" (see Fig. 6).

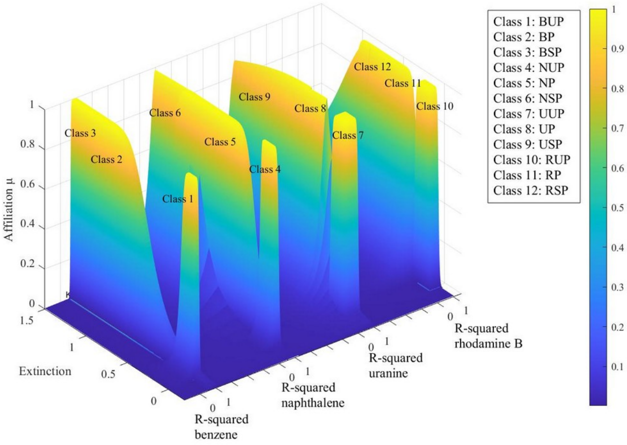


Fig. 6 Expert-based adaptation of the built classifier

In the ***working phase***, the classification model created in this way was used for the fuzzy identification of the current water samples. The work data are represented here by "artificial" work data, since they were generated by means of the original learning data set. This means that the total function was again overlayed on the spectra and the R-squared and the extinction maximum were determined. For a given substance or mixture of substances with the corresponding characteristic values for R-squared and extinction maxima, a characterization can now be carried out by determining the membership to the fuzzy groups described by the fuzzy pattern classifier. The selection of the group can typically be made according to the highest membership value. As an example, this is demonstrated for a total of six water samples or their object data sets, see Table 2 and Fig. 7.

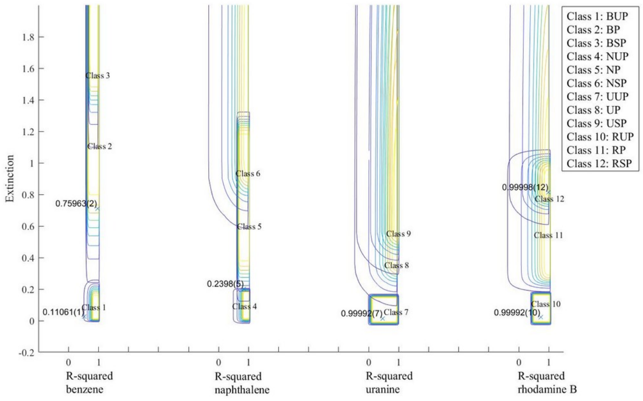


Fig. 7 Graphical representations of the assignment of test data

In Table 2, the largest value has been marked to illustrate the accuracy of the possible assignment. A partial superimposition of classes does not always allow a clear assignment. Nevertheless, interpretations can be made based on the calculated class membership.

Point 1 is only assigned to BUP with a very low membership value, which already indicates that benzene can only be present here with uncertainty. It can therefore be assumed that benzene is hardly present in this measurement. Point 2 is clearly assigned to BP with a value above 0.75. Point 6 is also clearly assigned to classes RP and RSP. Therefore, it can be concluded that benzene is present at point 2 and that rhodamine B is present for sure at point 6. At point 3, there is a low allocation to NP, whereby a tendency towards NUP is also recognisable. This fact can be explained by the same substance, but with different concentrations of these two substance mixtures, which is also reflected in the strong superposition of the corresponding classes. Here, the substance should be further observed to see in which direction it develops. In the case of points 4 and 5, both points are clearly allocated to a specific class with point 4 belonging to UUP and point 5 belonging to RUP. Since the concentrations here are very low, both are assigned to the classes that represent an uncertain presence of the respective substance.

IV. Conclusion

An approach to the characterization of water samples for on-site methods using fuzzy classification was presented. Measurement uncertainties during data acquisition and the

TABLE II.
MEMBERSHIP DATA OF THE TEST DATA FROM THE WORKING PHASE

No.	R-squared	Extinction-maxima	$\mu_{BUP,1}$	$\mu_{BP,2}$	$\mu_{BSP,3}$	$\mu_{NUP,4}$	$\mu_{NP,5}$	$\mu_{NSP,6}$	$\mu_{UUP,7}$	$\mu_{UP,8}$	$\mu_{USP,9}$	$\mu_{RUP,10}$	$\mu_{RP,11}$	$\mu_{RSP,12}$
1	0.5323070	0.0252742	0.1106125	0.010408	0.0003607	0.000563	0.000000	0.0044855	0.0000000	0.0019588	0.0018526	0.000000	0.0004723	0.00071276
2	0.9625610	0.7112262	0.0001511	0.759628	0.0297332	0.000000	0.000000	0.0072945	0.0000000	0.0021951	0.0021690	0.000000	0.0009161	0.00084021
3	0.8480867	0.2022218	0.0000000	0.000000	0.0000005	0.154456	0.239803	0.0227007	0.0000000	0.0081697	0.0071840	0.000000	0.0019905	0.00158935
4	0.4689550	0.0144004	0.0000000	0.000000	0.0000001	0.000000	0.0000000	0.0000006	0.9991880	0.0558700	0.0236306	0.000000	0.0007182	0.00320964
5	0.7336823	0.0212848	0.0000000	0.000000	0.00000000	0.000000	0.0000000	0.0000001	0.0000000	0.0000007	0.0000005	0.999922	0.0000001	0.0000000
6	0.9948569	0.8143327	0.0000000	0.0000000	0.0000000	0.0000000	0.0000000	0.0000001	0.0000000	0.0000006	0.0000004	0.0000000	0.9992831	0.9997920

associated fluctuations in the measured characteristic values can be modelled much more flexibly and more realistically than with conventional methods due to the fuzzy group description. The data necessary to obtain the membership functions can be obtained both by real measurements and by a linguistic description of different states by a human expert. Alternatively, a combined approach is possible. In summary, the consideration of uncertainties in the detection and evaluation of water samples is of great benefit. Firstly, the data can be modelled in a much more flexible and realistic manner by means of the implementation of fuzzy information. Secondly, data-based and/or expert-based modelling can be used (applicability to numerical or linguistic characteristics including mixed combinations of characteristics) which also offers an advantage for such methods. Finally, the modelling of states with fuzzy/incomplete description and the applicability to high-dimensional characteristic spaces can be realized.

Overall, the presented methodology offers a suitable approach for automatic classification of water sample data in on-site analysis. Successful field deployments for future applications require a more extensive data base with an increased number of characteristics for more detailed characterization of the water samples.

REFERENCES

[1] Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy (OJ L 327 22.12.2000 p. 1). (2006). In P. Sands & P. Galizzi (Eds.), Documents in European Community Environmental Law (pp. 879-969). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511610851.056.

[2] S. Zhuiykov, "Solid-state sensors monitoring parameters of water quality for the next generation of wireless sensor networks", Sens. Actuators B Chem., Bd. 161, Nr. 1, S. 1-20, 2012. DOI: https://doi.org/10.1016/j.snb.2011.10.078

[3] T. P. Lambrou, C. G. Panayiotou and C. C. Anastasiou, "A low-cost system for real time monitoring and assessment of potable water quality at consumer sites", Proc. IEEE Sensors, pp. 1-4, Oct. 2012. DOI: 10.1109/ICSENS.2012.6411190.

[4] T. P. Lambrou, C.C. Anastasiou, C. G. Panayiotou und M.M. Polycarpou, "A Low-Cost Sensor Network for Real-Time Monitoring and Contamination Detection in Drinking Water Distribution Systems", in IEEE Sensors Journal, Bd. 14, Nr. 8, S. 2765-2772, Aug. 2014. DOI: 10.1109/JSEN.2014.2316414.

[5] L. A. Zadeh, "Fuzzy Sets." - In: Information and Control 8, 338 – 353, 1965. DOI: https://doi.org/10.1016/S0019-9958(65)90241-X

[6] V. Traneva, S. Tranev and D. Mavrov" Interval-Valued Intuitionistic Fuzzy Decision-Making Method using Index Matrices and Application in Outsourcing" In: Proceedings of the 16th Conference on Computer Science and Intelligence Systems, M. Ganzha, L. Maciaszek, M. Paprzycki, D. Ślęzak (eds). ACSIS, Vol. 25, pages 251–254, 2021. DOI: http://dx.doi.org/10.15439/2021F77

[7] V. Traneva, S. Tranev, "Two-Stage Intuitionistic Fuzzy Transportation Problem through the Prism of Index Matrices" In: Position and Communication Papers of the 16th Conference on Computer Science and Intelligence Systems, M. Ganzha, L. Maciaszek, M. Paprzycki, D. Ślęzak (eds). ACSIS, Vol. 26, pages 89–96, 2021. DOI: http://dx.doi.org/10.15439/2021F76

[8] S. F. Bocklisch, “Prozessanalyse mit unscharfen Verfahren.” - Verlag Technik, Berlin, 1987. ISBN: 3-341-00211-1.

[9] V. Lohweg, C. Diederichs, D. Müller, “Algorithms for hardware-based pattern recognition.” EURASIP Journal on Applied Signal Processing 2004(12), 1912–1920, 2004. DOI: https://doi.org/10.1155/S1110865704404247

[10] U Mönks, V. Lohweg, and H. L. Larsen, “Aggregation Operator Based Fuzzy Pattern Classifier Design, Machine Learning in Real-Time Applications (MLRTA 09),” in KI 2009 Workshop, Paderborn | September 15th, 2009, accepted for Publication, 2009.

[11] S. F. Bocklisch, F. Bocklisch, M. Beggiato, J. F. Kremsa “Adaptive fuzzy pattern classification for the online detection of driver lane change intention, Neurocomputing,” Volume 262, 148-158, 2017. DOI: 10.1016/j.neucom.2017.02.089.

[12] F. Bocklisch, D. Hausmann, “Multidimensional Fuzzy Pattern Classifier Sequences for Medical Diagnostic Reasoning” In: Applied Soft Computing, Volume 66, 297-310, 2018. DOI: https://doi.org/10.1016/j.asoc.2018.02.041

[13] A.-J- Hempel, S. F. Bocklisch, “Parametric Fuzzy Modelling Framework for Complex Data-Inherent Structures” In: Proc. IFSA-EUSFLAT 2009, pp. 885–890, 2009. Corpus ID: 16501413.

[14] G. Herbst, “Online Recognition of Fuzzy Time Series Patterns” In: Proc. IFSA-EUSFLAT 2009. Corpus ID: 10151850.

[15] H. J. Zimmermann, “Fuzzy Set Theory and its Applications.” - Springer-Verlag, Berlin, 2001. ISBN: 978-94-015-7153-1.

[16] U. Scheunert, “Neue Verfahren der Fuzzy-Mengen-Verknüpfung und Fuzzy-Arithemtik und ihrer Anwendung bei der Sensor-Daten-Fusion. PhD thesis, TU Chemnitz, 2001.

Position Papers of the of the 17th Conference on Computer Science and Intelligence Systems pp. 35–39
DOI: 10.15439/2022F285
ISSN 2300-5963 ACSIS, Vol. 31

Detecting Uninformative Research Job Titles via Classifier Failures – Zero Shot Approach

Martin Víta
Faculty of Informatics and Statistics, Prague University of Economics and Business
Ekonomická 957, 148 00 Praha 4–Kunratice, Czech Republic
Email: info@martinvita.eu

***Abstract*—The aim of this paper is to introduce a novel approach to detecting "uninformative" job titles in research domain, i.e., detecting titles that convey little or no information about the focus and/or content of a particular job – like "Academic staff member AP/2", "PhD student position" etc. Such job titles decrease the success rate of job advertisements. The proposed approach belongs to zero shot approaches – it exploits only existing, easy accessible classification of jobs to research fields and it does not require any additional (manual) annotations. This work introduces an experimental corpus and provides preliminary results of our approach.**

I. Introduction

BASED on an internal survey of *ResearchJobs.cz*[1], job advertisements with well prepared, informative titles gain more attention from potential candidates in terms of (unique) users visits than vacancies having only general titles like "Postdoc position", "Academic staff member" etc. Moreover, as shown in [1], a job title is a suitable feature in predicting CTR[2] of job advertisements. Hence, a question of automated detecting of inappropriate job titles naturally arises in this setting.

Obviously, the task of detection uninformative titles can be straightforwardly addressed by common supervised ML techniques requiring an annotated corpus labeled in a binary way (informative/uninformative). However, a preparation of such a corpus is resource-extensive activity.

Our approach is based on the assumption that *an appropriate job title provides us enough information to classify the job advertisement to a correct research field.* Moreover, we assume that a classification of job offers to predefined fields is commonly available (usually selected from predefined categories by the user when submitting the advertisement).

Roughly said, *if a correctly working classifier of research fields assigns an incorrect label to a job title – since the label is known – then the job has an inappropriate title*, i.e., the failure of the classifier indicates an uninformative or even incorrect title. For example, if a classifier assigns "Computer science" label to a job entitled only "PhD student" submitted by the user within "Medical sciences" field, than we can conclude that the title is not appropriate, since it did not provide enough information to predict the research field correctly. In contrast, "Postdoc in therapy for neuromuscular diseases" labeled as "Medical sciences" by the user (who submitted the offer) and also by the classifier, than it indicates *sufficiently informative* title ensuring correct (automatic) classification.

Unlike ordinary classification task where *the text is the only input*, in our task, the input consists of the text (job title) as the first part and also of the human selected/submitted class as the second part. The result then depends on the *difference or identity of predicted and submitted class.*

Such a tool for detecting uninformative job titles can be directly used for an automatic feedback to users when submitting their advertisements and/or together with an automatic recommendation of a more suitable title.

The paper organization follows the standard IMRAD structure: Section 2 provides an overview about methods, i.e., models and data in our case. Section 3 contains results, Section 4 then the corresponding discussion. Since this paper has a "proof-of-concept character", the paper is completed with the overview of further work research directions.

II. Models and Data Involved

In this section, we provide a description of ML models and data used for training and zero shot task testing.

A. Core Idea of Zero Shot Approach

As already mentioned in Introduction, the keystone of the proposed approach is a classifier assigning a research field to a job title.

The trained classifier will be subsequently used to classify job titles from the (zero shot) test set where research field labels are known and these items are also equipped with a binary (informative–uninformative) human labels which serves as a gold-standard. If there is a mismatch of "real" research field label and the output of the classifier, than the title is marked as `uninformative`, otherwise marked as `informative`. The evaluation w.r.t. these two labels is performed further in a standard way.

The basic dataset to be used in this work (see Subsection 2.3) contains 5,341 positions from Euraxess portal, thus the same number of job titles. The median length (number of characters) of job titles is 57, whereas the the average is 65.76, 1st quartile 36 and 3rd quartile 85 characters. Thus we are dealing with classification of short texts.

[1] Czech job portal focused on research and academic vacancies

[2] Click-Through Rate, CTR is defined as the number of clicks that a given advertisement receives divided by the number of times it is shown.

Classification of short texts belongs to one of the traditional tasks of ML/NLP with a long history [2]. This direction of research was often driven by motivation for sentiment analysis of tweets [3] and other social media content.

B. Models for Classification

The general task of classification of short text can be tackled by several ML methods. However, the aim of this work is not to focus on the classification itself – but later on the zero shot [4] part. This section provides an overview of models as well as corresponding features involved. In this work we will deal only with neural networks based models, namely:

- Character-based 1D-convolutional neural network (CharCNNs),
- Convolutional Word2Vec-based model (CNNs),
- Universal sentence embeddings (USE).

CharCNN Character-based convolutional networks are a frequent choice for processing of short texts [5]. Their advantages are – among others – that they can be employed in language agnostic setting [3], they are robust to misspellings and they can easily deal with special character combinations such as emoticons etc.

In this model, a job title is represented as a fixed length sequence (255 characters, since it is the maximal length of the job title) of one-hot encoded character vectors – in this case we deal with 128 characters, shorter titles are padded with zero vectors. Therefore, the corresponding matrix has dimension 128×255. This matrix is subsequently processed by 1D-convolutional layer with 25 filters of kernel size equal to 3 (i.e., we are processing "character-trigrams") and the result of this convolutional layer is fed to 1d-max pooling layer with pool-size again equal to 3. The decision is made by standard softmax dense layer with 9 output neurons (i.e., number of output classes). The final model contains 28,759 trainable parameters, the hyperparameters of the model (number of filters etc. were set using grid search).

Diagram of the architecture is shown on Figure 1[3].

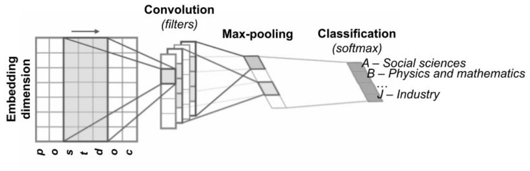


Fig. 1. Character-based CNN architecture

CNNs Convolutional word2vec-based models belong to traditional architectures for text classification [2]. Convolutional neural networks were successfully used in short text classification (tweets in particular) in many branches, including biomedical domain [6].

In this setting, a job title is represented of a fixed length (30 words in our case) sequence of word2vec [7] embeddings. Shorter titles are padded by zero vectors. Since we deal with texts of research domain, we did not used general word2vec embeddings but pretrained embeddings of dimension 200 learned on texts of scientific (biomed) domain, that were used in [8]. Representation of words that occur in a job title but are not contained among words with pretrained embeddings is uniformly set to zero vectors.

This sentence matrix (200×30) is fed to a 1D-convolutional layers with 22 filters and kernel size 3 followed by 1d-max-pooling layer with pool-size of 3. Hyperparameters were set again using grid search. Finally, softmax classification output layer is used. This model has 15,211 trainable parameters in total, the overall architecture is depicted on Figure 2 and it is formally similar to the previous case, however, here we do not use one hot encoding.

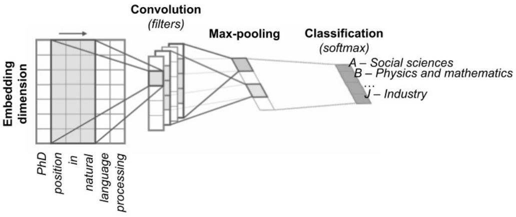


Fig. 2. Word2vec-based CNN architecture

USE As an example of more advanced methods – transformer-based representations, we used *Universal Sentence Encoder (USE)* [9], successfully applied in many areas such as Semantic Textual Similarity (STS), [10]. The trained model implementation was obtained from TensorFlow Hub[4]. It provides a 512-dim sentence (text snippet) representations – in our work, the pretrained network was used straightforwardly for feature extraction. These representations were subsequently fed into a dense layer (dim: 32; the number was obtained again by grid search), followed by the output softmax layer as in the previous cases. The model has 16,713 trainable parameters in total.

Implementation Details: The complete implementation of this work was elaborated in R + Keras library[5]. As optimizer, RMSprop [11] was used in all training scenarios. The number of epochs varies from 12 to 16 depending on the model and data involved.

C. Data Involved

The data used in this work can be basically divided into two groups: data used for training the "research field classifier" and data used for testing the zero shot approach (informative/uninformative classification).

[3] Diagrams are modified versions of one from: https://towardsdatascience.com/convolutional-neural-network-in-natural-language-processing-96d67f91275c

[4] https://tfhub.dev/google/universal-sentence-encoder/4

[5] https://cran.r-project.org/web/packages/keras/index.html

TABLE I
DISTRIBUTION OF LABELS IN BASIC JOB TITLES DATASET

Research Field (class)	Instances
A – Social sciences	754
B – Physics and Mathematics	541
C – Chemistry	224
D – Geosciences	117
E – Biosciences	309
F – Medical Sciences	515
G – Agriculture	461
I – Informatics/computer science	262
J – Industry	588

1) Data for Learning the "Research Field Classifier": The key dataset is a database dump of Euraxess portal[6] from March, 2021. It was provided in the form of one large XML file. Each position has several attributes, however, from our point of view, only a few of them are relevant: job title and research field, in particular.

The raw dataset contains 5,341 positions. Each position is assigned to at least one research field (for example: *Psychological sciences*, *Physics* etc.) and may be assigned also to research subfields (for example: *Psychology*, *Applied physics* etc.). The total number of research fields is 41, including two special labels *All* and *Other*.

To ensure to deal with a single-class classification task, we filtered out only positions that have just only one research field label, and moreover, we did not take into account positions having *All* or *Other* labels. This resulted in a reduced dataset of 3,771 positions whereas each position is labeled by one of 39 labels. However, this labeling is strongly unbalanced – top 3 classes are *Engineering*, *Agricultural sciences* and *Medical sciences* containing 465, 461 and 446 positions respectively. On the other hand, the least numerous labels in this reduced dataset are *Criminology*, *Ethics in social sciences* and *Ethics in physical sciences* with 1, 1 and 2 occurrences respectively.

In order to deal with more balanced classification and to reduce the number of classes, we used a coarser Czech classification system of research branches[7] having 9 classes (more precisely, it deals with 10 classes, the last one is K – Defense, but this field is not taken into account). Simple handcrafted transformation rules were prepared. The utilization of this classification has also other reasons that will be obvious later in this chapter. The distribution of labels in the dataset of position titles is provided in Table I.

Subsequently, we randomly selected 3,000 of items (positions) to be the training set for research filed classification. The rest of 771 positions were left for further preparation of test set of zero shot ("informative/uninformative" classification task).

To achieve better classification accuracy, we also prepared an auxiliary annotated dataset of a bigger volume – research project titles together with their research branch classification. These data were taken from open data section of Czech R&D Information System[8] which gathers (meta)data about all R&D projects in the Czech Republic funded by public sources. This dataset contains 43,694 items (project name–classification pairs). The aim of exploiting this dataset was to extend the original training data by this easily obtainable stuff. Each model is trained both with the original basic dataset and this enriched one.

To provide a better idea of items in this auxiliary dataset, we randomly select three examples of project titles with corresponding classification labels.

- *Phospholipid metabolizing enzymes as new components of salicylic acid signalling pathway:* C – Chemistry
- *Communities and resources in late prehistory of Jebel Sabaloka, central Sudan: from analysis to synthesis:* A – Social sciences
- *Optimization of hunted species management in relation to the sustainable forest management:* G – Agriculture

2) Test Data for Zero Shot Classification: The second part of the data involved is the test dataset for zero shot classification, i.e., job titles manually labeled as informative or uninformative.

There were 771 remaining jobs (job titles) from Euraxess dataset that were not intended for training, whereas 102 (!) of them were manually marked as uninformative; the rest (i.e., 669 items) is considered as informative. This dataset of 102 uninformative job titles was subsequently enriched by another set of 48 uninformative titles (annotated manually again) which were obtained from a randomly shuffled collection of jobs from ResearchJobs.cz portal (this portal uses also the "A–J research branches" classification). Hence the number of uninformative examples in the test dataset reached 150. To obtain a balanced test set, 150 items with *informative* titles were randomly selected from already mentioned list of 669 items. This dataset can be provided upon a (mail) request. The content of uninformative subset of job titles is illustrated using a wordcloud, see Figure 3. Inter-annotator agreement was not investigated in this context.

Obviously, typical, i.e., most frequent, words in uninformative part of job titles are general names of academic/research positions (professor, PhD), words linked to hiring process (call, applications), general duties (teaching, research).

In addition to general words common for both classes, the informative job titles contain bigger amount of relatively infrequent words denoting particular research fields (physics, biology) and corresponding specific words (quantum, molecular).

Examples of informative and uninformative job titles randomly selected from this zero shot test dataset are provided in the following list (job title – user selected research field classification – informative/uninformative label):

- Doctoral student in Economic History (A – Social sciences): informative

[6] Euraxess portal is one of the most important European job portals focused on research and academic position. It publishes positions solely in English – https://euraxess.ec.europa.eu/jobs/.

[7] IS VaVaI: https://www.isvavai.cz/

[8] https://www.isvavai.cz/open-data

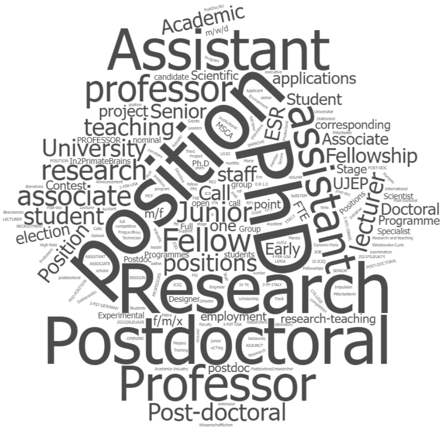


Fig. 3. Wordcloud generated from uninformative job titles

TABLE II
RESULTS IN DIFFERENT SCENARIOS (MODEL–DATA)

Architecture	Dataset	10-fold	UninfTask
CharCNN	basic	0.4413	**0.6900**
	ext	0.4100	0.6167
CNN	basic	0.5717	0.7600
	ext	0.5720	**0.7833**
USE	basic	0.5480	***0.7900***
	ext	0.5897	0.7867

- PhD scholarship in 6G Wireless Communications (J – Industry): informative
- Assistant Professor FSI UJEP (J – industry): uninformative
- Fellowship for Postdoctoral Researcher (F – Medical sciences): uninformative

III. RESULTS

Since our work relies on two-way classification, the evaluation is based mainly on accuracies. The evaluation has basically two levels: evaluation of research field classifier and evaluation of uninformative/informative classifier.

The evaluation of research field classifier is done as an average of accuracies in 10-fold cross validation, the evaluation of uninformative/informative classifier as standard accuracy.

The results in different scenarios (model-data) are summarized in Table II

Confusion Matrices: More detailed view on bold-marked results (i.e., best results for each model) are available via confusion matrices: Table III, Table IV and Table V.

TABLE III
CONFUSION TABLE FOR USE MODEL

		Predicted label	
		Uninformative	Informative
True label	Uninformative	126	24
	Informative	39	111

TABLE IV
CONFUSION TABLE FOR CNN MODEL

		Predicted label	
		Uninformative	Informative
True label	Uninformative	122	28
	Informative	37	113

IV. DISCUSSION

The best results were achieved using Universal Sentence Encoder. As can be seen from the confusion matrix, our algorithm based on research field classifier failures was able to detect 79 % of uninformative job titles. It should be mentioned that this proposed approach inherently implies certain error arising from the fact that in some cases the classifier can predict the correct class of uninformative title by chance.

On the other hand, 26 % of informative job titles were marked as uninformative – i.e., predicted and true label were *not equal* in the case of informative title (“informativeness” was labeled manually). Preliminary human conducted analysis indicates that most of these cases were borderline items with respect to output classes (classification system) and the the assumption of dealing with jobs that are *assigned just to one class* in our setting. A position “PhD in robotics” can serve as an example: both *Computer science* and *Industry* labels are relevant in this case, analogous situation is often between medical and biological sciences – the correct “real-world” assignment is the subject of the *whole text of job detail* which is not taken into account due to the main aim of this work. Hence an alert when predicted research field label and label selected by the user are not identical as a side-effect points out possibly confusing title.

According to observations of confusion matrix of CNN approach, we see that both USE and CNN are comparable. In the number of false negatives, CNN approach slightly outperforms USE, in true positives, the situation is reversed. The results of CharCNN are strongly below expectations.

In both successful approaches (USE and CNN) the effect of additional training items (research project titles) is marginal, moreover, in CNN approach training without additional data lead to better performance. Notable effect of enriching the training dataset was observed only in convolutional character-based approach.

The cases of *trully informative job titles labeled as uninformative* will be a subject of further investigations on a larger

TABLE V
CONFUSION TABLE FOR CHARCNN MODEL

		Predicted label	
		Uninformative	Informative
True label	Uninformative	123	27
	Informative	66	84

dataset. Generally, the sources of this misclassification belong to the following two groups:

1) wrong category assignment by user during submission of the job advertisement,
2) incorrect work of the research field classifier.

Relatively poor performance of the research field classifier in 10-fold cross validation at first sight is caused by the following reasons:

1) High proportion of uninformative job titles in the Euraxess dataset: according to our preliminary human experiments it consists approximately 1/8 of the dataset.
2) Frequent presence of borderline case (as described above).
3) Relatively high number of output classes (thus also a trivial majority vote classifier achieves very low accuracy).
4) Occasional occurrence of job titles in languages other than English, misspellings etc.

V. Conclusion and Further Work

We have introduced a novel zero shot approach to detection of uninformative job titles in research domain based on exploiting incorrect predictions of a job field classifier. We prepared corresponding experimental corpora and provide some preliminary results.

Further Work

Our results of this zero shot approach indicate that this chosen direction is promising. Nevertheless, there is a large room for improvement, mainly in the sense of exploiting fine-tuned variants of BERT [12] and its variants like SentenceBERT [13] and others [14].

For our preliminary experiments, the single label setting was chosen due its simplicity. However, the nature of the task is rather multi-label, thus we will adopt our approach for multi-label classification. The side effect is that we can immediately use larger datasets (without filtering jobs that are assigned just to one class).

Further generalization may lead also to fuzzy point of view: rather than speaking about crisp "text–class" membership function we can deal with a fuzzy membership: each job advertisement (and job title so) may belong to more classes with different degrees of membership – as an example we can consider positions like "Postdoc in cancer research" which are spanned between biological and medical sciences. Fuzzy approach to sentiment analysis of tweets [15] can serve as an inspiration.

As already mentioned, this work is restricted only to job titles in English. Another direction of further investigations can be naturally focused on language agnostic as well as multilingual approaches (analogous to [16] for instance) which will be able to detect uninformative titles also in other languages.

A separate chapter in further research is a language generation for improving job titles – given a text, i.e., content of the job advertisement, the task is to create an appropriate, *informative* job title. As a promising direction seems to be application of GPT transformers [17] for language generation as well as summarization techniques [18] – extreme summarization of particular parts of job detail (e.g., requirements) may be a suitable addition to an uninformative prefix (like "Postdoc", "PhD student" or "Assistant Professor").

References

[1] M. Jiang, Y. Fang, H. Xie, J. Chong, and M. Meng, "User click prediction for personalized job recommendation," *World Wide Web*, vol. 22, no. 1, pp. 325–345, 2019. doi: 10.1007/s11280-018-0568-z

[2] K. Kowsari, K. Jafari Meimandi, M. Heidarysafa, S. Mendu, L. Barnes, and D. Brown, "Text classification algorithms: A survey," *Information*, vol. 10, no. 4, p. 150, 2019. doi: 10.3390/info10040150

[3] J. Wehrmann, W. Becker, H. E. Cagnini, and R. C. Barros, "A character-based convolutional neural network for language-agnostic twitter sentiment analysis," in *2017 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2017. doi: 10.1109/IJCNN.2017.7966145 pp. 2384–2391.

[4] P. K. Pushp and M. M. Srivastava, "Train once, test anywhere: Zero-shot learning for text classification," *arXiv preprint arXiv:1712.05972*, 2017. doi: 10.48550/arXiv.1712.05972

[5] X. Zhang, J. Zhao, and Y. LeCun, "Character-level convolutional networks for text classification," *Advances in Neural Information Processing Systems*, vol. 28, 2015.

[6] L. Akhtyamova, M. Alexandrov, and J. Cardiff, "Adverse drug extraction in twitter data using convolutional neural network," in *2017 28th International Workshop on Database and Expert Systems Applications (DEXA)*. IEEE, 2017. doi: 10.1109/DEXA.2017.34 pp. 88–92.

[7] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," *Advances in Neural Information Processing Systems*, vol. 26, 2013.

[8] G.-I. Brokos, P. Malakasiotis, and I. Androutsopoulos, "Using centroids of word embeddings and word mover's distance for biomedical document retrieval in question answering," in *Proceedings of the 15th Workshop on Biomedical Natural Language Processing*, 2016. doi: 10.18653/v1/W16-2915 pp. 114–118.

[9] D. Cer, Y. Yang, S.-y. Kong, N. Hua, N. Limtiaco, R. S. John, N. Constant, M. Guajardo-Cespedes, S. Yuan, C. Tar *et al.*, "Universal sentence encoder," *arXiv preprint arXiv:1803.11175*, 2018. doi: 10.48550/arXiv.1803.11175

[10] D. Cer, M. Diab, E. Agirre, I. Lopez-Gazpio, and L. Specia, "Semeval-2017 task 1: Semantic textual similarity - multilingual and cross-lingual focused evaluation," *arXiv preprint arXiv:1708.00055*, 2017. doi: 10.48550/arXiv.1708.00055

[11] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT press, 2016.

[12] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018. doi: 10.48550/arXiv.1810.04805

[13] N. Reimers and I. Gurevych, "Sentence-bert: Sentence embeddings using siamese bert-networks," *arXiv preprint arXiv:1908.10084*, 2019. doi: 10.48550/arXiv.1908.10084

[14] A. Gillioz, J. Casas, E. Mugellini, and O. Abou Khaled, "Overview of the transformer-based models for nlp tasks," in *2020 15th Conference on Computer Science and Information Systems (FedCSIS)*. IEEE, 2020. doi: 10.15439/2020F20 pp. 179–183.

[15] C. Jefferson, H. Liu, and M. Cocea, "Fuzzy approach for sentiment analysis," in *2017 IEEE international conference on fuzzy systems (FUZZ-IEEE)*. IEEE, 2017. doi: 10.1109/FUZZ-IEEE.2017.8015577 pp. 1–6.

[16] A. F. M. de Paula, R. F. da Silva, and I. B. Schlicht, "Sexism prediction in spanish and english tweets using monolingual and multilingual bert and ensemble models," *arXiv preprint arXiv:2111.04551*, 2021. doi: doi.org/10.48550/arXiv.2111.04551

[17] B. Ghojogh and A. Ghodsi, "Attention mechanism, transformers, bert, and gpt: Tutorial and survey," 2020.

[18] I. Cachola, K. Lo, A. Cohan, and D. S. Weld, "Tldr: Extreme summarization of scientific documents," *arXiv preprint arXiv:2004.15011*, 2020. doi: 10.48550/arXiv.2004.15011

4th International Workshop on Artificial Intelligence in Machine Vision and Graphics

THE main objective of the 4th Workshop on Artificial Intelligence in Machine Vision and Graphics (AIMaViG'22) is to provide an interdisciplinary forum for researchers and developers to present and discuss the latest advances of artificial intelligence in the context of machine vision and computer graphics. Recent advancements in artificial intelligence resulted in the rapid growth of both methods and applications of machine learning approaches in computer vision, image processing, and analysis. The development of parallel computing capabilities in the first decade of the 21st century that boosted the development of deep neural networks became a real gamechanger in machine vision. The workshop covers the whole range of AI-based theories, methods, algorithms, technologies, and systems for diversified and heterogeneous areas related to digital images and computer graphics.

Topics

The topics and areas include but are not limited to:

- image processing and analysis:
  - image enhancement,
  - linear and non-linear filtering,
  - object detection and segmentation,
  - shape analysis,
  - scene analysis and modeling,
  - scene understanding,
- machine learning for vision and graphics:
  - pattern recognition,
  - deep neural models,
  - convolutional networks,
  - recurrent networks,
  - graph networks,
  - generative adversarial networks,
  - neural style transfer,
  - deep reinforcement learning,
- machine vision:
  - image acquisition,
  - stereo and multispectral imaging,
  - embedded vision,
  - robotic vision,
- image theory:
  - computational geometry,
  - image models and transforms,
  - modeling of human visual perception,
  - visual knowledge representation and reasoning,
- visualization and computer graphics:
  - data-driven image synthesis,
  - graphical data presentation,
  - computer-aided graphic arts and animation,
- applications:
  - innovative uses of graphic and vision systems,
  - image retrieval,
  - autonomous driving systems,
  - remote sensing,
  - digital microscopy,
  - security and surveyance systems,
  - document analysis,
  - OCR systems.

Technical Session Chairs

- **Iwanowski, Marcin,** Warsaw University of Technology, Poland
- **Kwaśnicka, Halina,** Wrocław University of Science and Technology, Poland
- **Śluzek, Andrzej,** Khalifa University, United Arab Emirates

Program Committee

- **Andrysiak, Tomasz,** Bydgoszcz University of Science and Technology, Poland
- **Angulo, Jesús,** Mines ParisTech, France
- **Cyganek, Bogusław,** AGH University of Science and Technology, Poland
- **Kasprzak, Włodzimierz,** Warsaw University of Technology, Poland
- **Kwolek, Bogdan,** AGH University of Science and Technology, Poland
- **Okarma, Krzysztof,** West Pomeranian University of Technology, Poland
- **Olszewski, Dominik,** Warsaw University of Technology, Poland
- **Palus, Henryk,** Silesian University of Technology, Poland
- **Subbotin, Sergey,** Zaporozhye National Technical University, Ukraine
- **Tomczyk, Arkadiusz,** Lodz University of Technology, Poland

DOI: 10.15439/2022F283
ISSN 2300-5963 ACSIS, Vol. 31

New Thermal Automotive Dataset for Object Detection

Tomasz Balon*, Mateusz Knapik†‡ and Bogusław Cyganek†
*Department of Electrical Engineering, Automatics, Computer Science and Biomedical Engineering,
Email: *tbalon@student.agh.edu.pl*
†Department of Computer Science, Electronics and Telecommunication
Email: *mknapik@agh.edu.pl, cyganek@agh.edu.pl*
AGH University of Science and Technology,
Al. Mickiewicza 30, 30–059 Kraków, Poland
‡MyLED Inc.
Email: *m.knapik@myled.pl*
Ul. W. Łokietka 14/2, 30–016 Kraków, Poland

Abstract—**Although there are many efficient deep learning methods, object detection and classification in visible spectrum have many limitations especially in case of poor light conditions. To fill this gap, we created a novel thermal video database containing few thousands of frames with annotated objects acquired in far infrared thermal spectrum. Thanks to this we were able to show its usability in the traffic object recognition based on the YOLOv5 network, properly trained to gain maximal performance on thermal images, which contain many small objects and are characteristic of different properties than the visible spectrum counterparts. The proposed thermal database, as well as the fully trained model are main contributions of this paper. These are made available free for other researchers. Additionally, based on the highly efficient car detector we show its application in the car speed measurement based exclusively on thermal images. The proposed system can be also used in the Advanced Driver-Assistance Systems (ADAS), and help autonomous driving.**

I. Introduction

ARTIFICIAL intelligence (AI) and machine learning (ML) are two of the fastest developing technologies nowadays. New and novel architectures are developed to be faster, more accurate and reliable. Image classification and object detection is very active field of research and many innovative techniques were proposed recently. Range of possible applications is very wide and autonomous driving is one of them. It gained much of an interest from scientists and companies recently. Vision systems based on a visible lights have limitations when used in a moving vehicle caused by wide range of lighting conditions that can occur. Low light during the night time as well as very high amounts of light during the day pose a challenge to hardware and software modules. On the other hand, thermal imaging in recent years gains popularity, both in industrial solutions, as well as in research projects.

However, the development of the image analysis methods might be rapid for images acquired using conventional RGB sensors, other imaging technologies operating in spectrum beyond visible light still fall short mostly due to the lack of publicly available sufficiently large training datasets.

To help alleviate this problem, in this paper we present:

- A new novel traffic dataset acquired using thermal imaging camera.
- Pretrained object detection model based on YOLOv5 architecture.
- An exemplary application based on detections: speed measurement in thermal spectrum.
- Examples of potential further applications.

Our dataset contains videos with close to 30,000 hand-annotated objects, many of small size, which makes them difficult to detect. Our second contribution is pretrained object detection model based on YOLOv5 architecture. Primary use case intended for this model is detecting four classes of objects in thermal images, as well as car speed measurement, which is the third contribution provided in this paper.

The paper is organized as follows. Section II describes the related works. In Section III process of acquiring the data and model training is explained. Section III-B presents the structure and provides more insight into dataset. Section IV shows example of how acquired data might be used in calculating vehicle's speed. In Section V more of future development possibilities are discussed. Finally, Section VI concludes the paper.

II. Related works

Object detection combined with thermal imaging gained a lot attention in recent years. Thermal imaging is based on observing infrared waves emitted by warm objects [1]. It allows user to see infrared spectrum which is invisible with naked eye. Hence it's willingly used not only during daylight, but especially during nighttime or difficult weather conditions [2] [3]. In this section an overview of the influencing works related to the processing of thermal images, analysis and detection in infrared spectrum is presented and discussed.

A. Object detection in thermal images

Knapik et al. [4] presented eye detection in thermal images scheme using the virtual high dynamic range technique, to enhance performance of the dense grid of scale-invariant feature descriptors, combined with the bag-of-visual-words approach.

Redmon et al. proposed a series of improved versions of the YOLO architecture, i.e. YOLOv2 [5] and YOLOv3 [6]. Deep convolutional backend network, along with techniques like residual skip connections, residual blocks and upsampling, it is still one of the fastest object detection techniques, while achieving very respectable accuracy. Recently, a thorough refreshment of the YOLO architecture, named YOLOv5, was presented by Jocher et al. [7].

In [8] Bhattarai and MartíNez-Ramón presented intelligent system for real-time object detection and recognition for firefighters during an emergency response. They trained deep Convolutional Neural Network (CNN) to improve situational awareness by identifying objects of interest from thermal imagery in real-time.

In an article from 2020, Gong et al. [9] employed thermal camera for vehicle detection task. In order to achieve faster detection time, the modified YOLOv3-tiny architecture, by recalculating anchor box priors as well as deepening the network structure.

Thermal images are also used to enhance other modalities. Zhou et al. in [10] presented feature fusion network for salient object detection (SOD) task, merging foreground and background information from RGB camera and thermal sensor. Proposed architecture outperforms 12 state-of-the-art methods under different evaluation indicators.

Some researchers propose custom network architectures, designed specifically for infrared images, like Dai et al. in [11]. They proposed TIRNet architecture, which consist of lightweight feature extractor as well as residual branch for regression and classification.

B. Thermal imaging and datasets

One of the biggest problems of thermal imaging is low resolution. To mitigate this problem, Rivadeneira et al. presented novel super-resolution architecture for thermal images based on CycleGAN network [12]. Authors created they own dataset for network training.

Yeduri et al. presented novel low resolution thermal images dataset in [13]. Containing 3200 images of sign language digits captured with very low-res thermal sensor can be used to build human input devices for people with disabilities.

Kristo et al. [2] compared night vision to thermal imaging in their paper and emphasized benefits of using infrared thermal imaging approach over standard RGB. Their research was focused on difficult weather conditions, as described, their dataset was captured during winter in different weather conditions, such as rain, fog or clear weather, during the night. In their paper, YOLOv3 model was trained on custom dataset to detect objects (people) even from far away (up to 215m).

System proposed by Knapik and Cyganek in [14] proved that thermal imaging can be successfully applied to driver's fatigue detection task based on yawn detection. Face alignment is done by detection of eye corners. Then, yawns are detected based on the proposed yawning thermal model.

Thermal imaging was proposed by Farooq et al. [15] to support Advanced Driver-Assistance Systems (ADAS). In their research, thermal imaging was used to capture different objects that are likely to be met on road, such as person, dog, bicycle, car, bike, etc. They also proposed YOLOv5 model trained on their custom dataset. In their work, also a comparison between several available YOLOv5 models was made.

III. Experimental part

A. Data acquisition

The data provided with this article was collected using the thermal imaging camera FLIR® A35. Acquisition took place in the afternoon, between 3.30 PM and 4 PM with cloudy weather and temperature around -3°C. Videos contain real-life traffic with cars, trucks, buses and people. Camera was placed at elevated footbridge above the street, our setup is shown in Figure 1/ Figure 2 shows RGB image of field of view the setup had. It also depicts exact weather conditions and approximate time of the day.

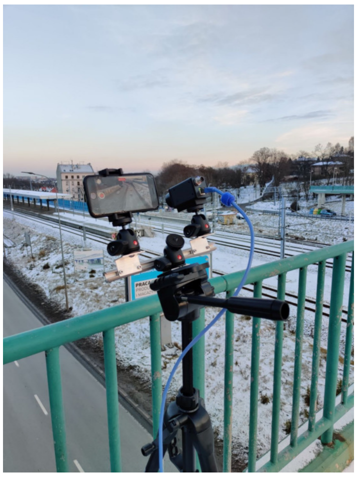


Figure 1: Acquisition setup

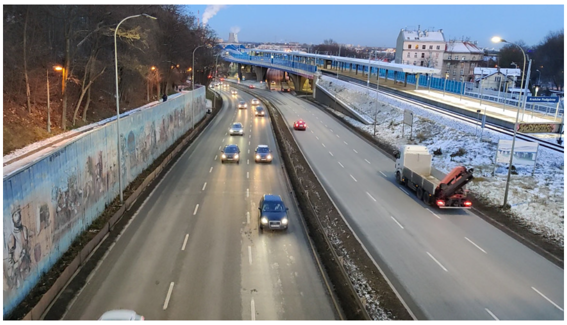
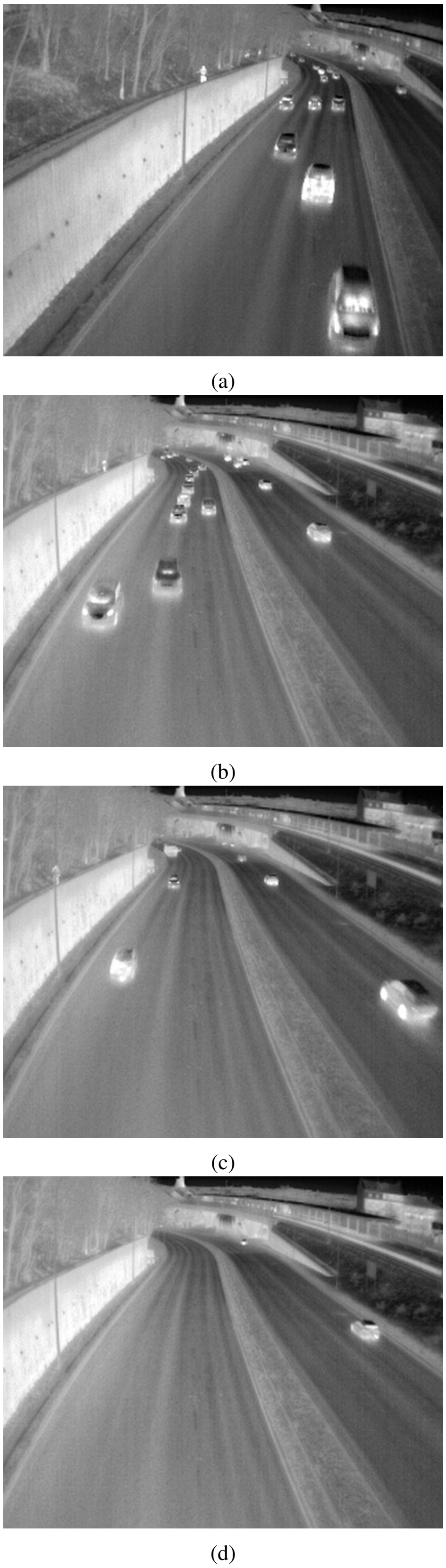


Figure 2: Camera's field of view

Images were manually labeled to provide ground-truth data which is used for training and evaluation. Sample images from the dataset are presented in Figure 3.


(a)

(b)

(c)

(d)

Figure 3: Sample images from the dataset

B. Dataset description

Dataset contains over 6000 annotated images with more than 30000 object instances. Within dataset, 4 classes are annotated, as shown in Table I. Frames were annotated using DarkLabel [16] software. To maintain consistency with COCO dataset, we used the same class IDs. Figure 4 presents number of the instances per class.

Table I: Class IDs

Number	Name
0	person
2	car
5	bus
7	truck

Images in the dataset are in .jpg and .bmp format. The resolution of single image is 320x256 pixels with 8-bit grayscale values. Annotation files are stored in text files with .txt extension in YOLO format [7].

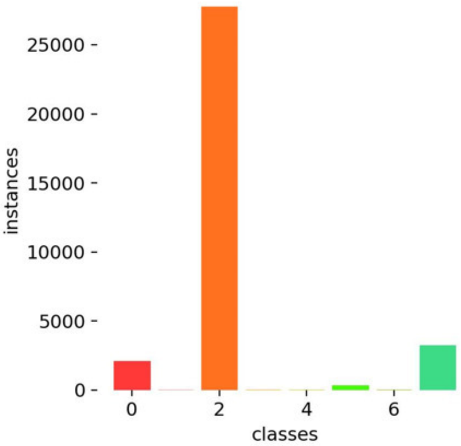


Figure 4: Number of instances in each class.
Classes: 0 - person, 2 - car, 5 - bus, 7 - truck

Dataset is publicly available for all researchers to download from our website: https://home.agh.edu.pl/~cyganek/AutomotiveThermo.zip.

C. Data structure

Dateset contains images and labels as well as trained YOLOv5 object detector. Images are divided into train, validate and test subsets, each stored in a separate folders.

D. Object detection model training

To evaluate the dataset, we decided to train and test object detection model based on YOLO architecture. We chose open-source implementation provided by Ultralytics company [7]. This architecture was chosen due to its availability and ease of use and high quality of code. Due to dataset size and computation speed we decided to use YOLOv5m variant out of other available models (Figure 5). Training was executed in several runs, each with slightly different variables, such as e.g. epochs number. Finally, after comparing results of all runs, model was trained for 50 epochs, to observe if this would lead to model overfitting. Training results are presented in Figure 6. These results contain graphs of loss subfunctions: box loss, objects loss and classification loss as well as precision, recall and mean average precision (mAP). It is clearly visible, that the more epochs pass by, the more accurate the model is. Figures 7a - 7d show precision, recall and overall score of the model's accuracy. After being trained, model was later tested on new, unseen but labeled images. Figure 8 shows predictions of labeled objects on test data. A closer look of a predictions made by trained model and the confidence levels of detected objects are shown in Figure 9.

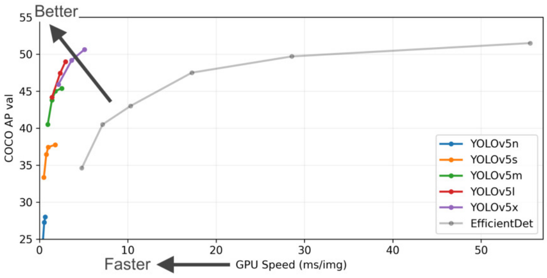


Figure 5: Family of YOLOv5 models
Source: [7]

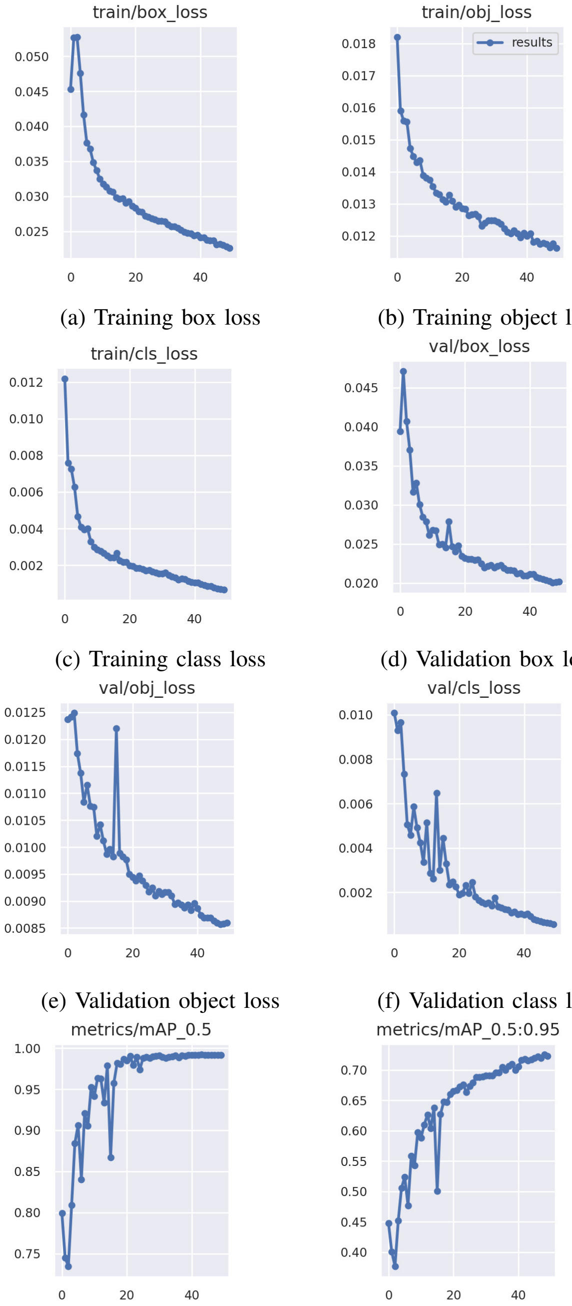
IV. Vehicle speed measurement

Presented dataset might be used for vehicle speed measurement. This can be achieved by calculating distance travelled by a vehicle within some portion of time. It's relatively easy to get the timestamps, as images come with exact date with minutes and seconds in their name. When it comes to getting distance out of collected data, let's take into consideration 2 photos (Figure 10). Timestamp provided with left image is 15:50:44, and timestamp provided with right image is 15:50:45, which means, that exactly one second passed between taking those two images. Let's also consider car marked in yellow circle.

Now, to measure the distance the vehicle has travelled within this time, we need to have some reference. This can be done in several ways, but for our sample application we decided to use the line marks between right and middle lane. Although they are not clearly visible, they still can act as a reference in this experiment. Based on knowledge where the recordings took place and the standards according to which the stripes are painted [17], we can conclude that stripe itself is 2m long, and the gap between 2 stripes is 4m long.

In Figure 11 red lines show there the stripes are, and blue lines represent car's front in regard to the stripes. Now, distance can finally be measured. In Figure 11, considered vehicle travelled 3 gaps and 2.5 stripes, which is equal to

$$3 \cdot 4m + 2.5 \cdot 2m = 17m \tag{1}$$



(a) Training box loss

(b) Training object loss

(c) Training class loss

(d) Validation box loss

(e) Validation object loss

(f) Validation class loss

(g) Metrics mAP_0.5

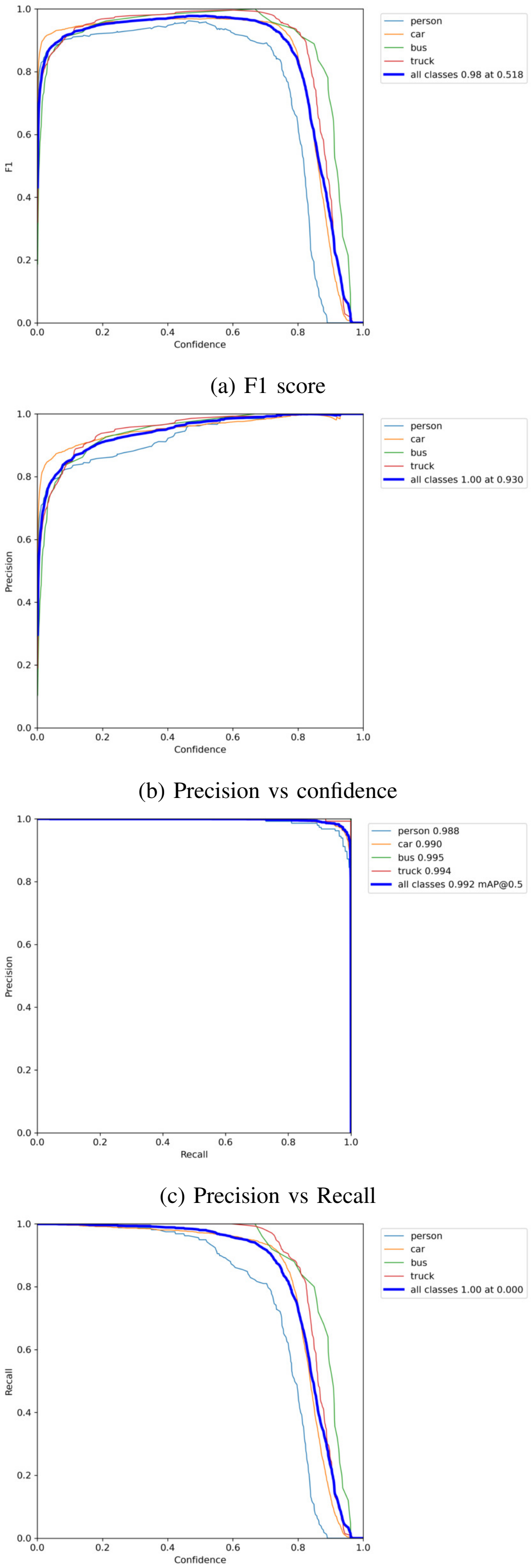
(h) Metrics mAP_0.5:0.95

Figure 6: YOLO model train results

All necessary data to calculate the velocity is now available, thus

$$\frac{17m}{1s} \cdot \frac{3600\frac{s}{h}}{1000\frac{m}{km}} = 61.2\frac{km}{h} \tag{2}$$

Although no radar data was acquired to back up this calculation, the result is believable and within allowed speed limit on this road, which leads to conclusion, that velocity can be measured by calculating distance travelled within certain timestamped frames.



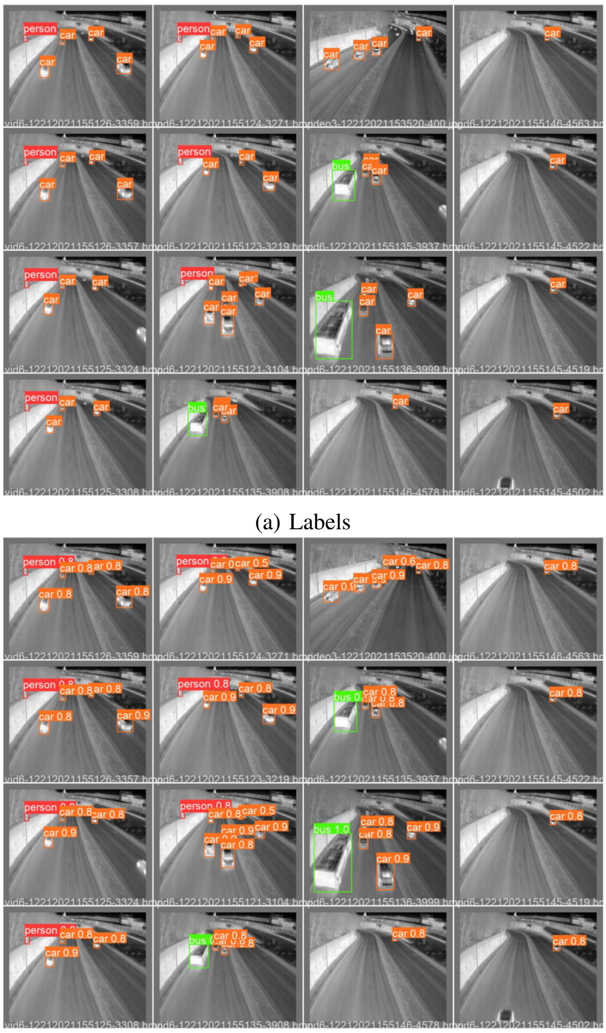
(a) F1 score

(b) Precision vs confidence

(c) Precision vs Recall

(d) Recall vs confidence

Figure 7: Training results.



(a) Labels

(b) Predictions

Figure 8: Detection results on test data

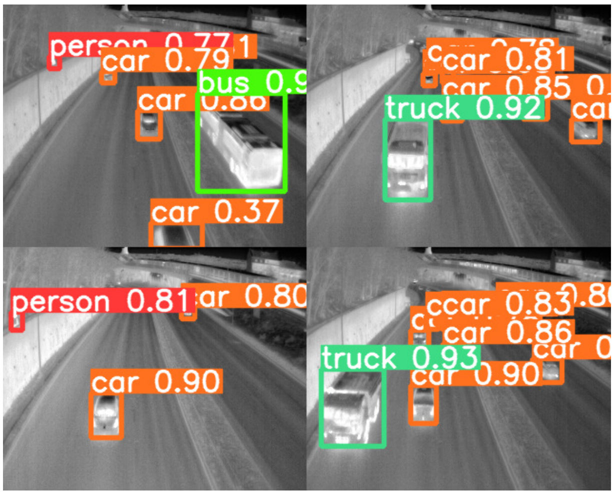


Figure 9: Sample detections

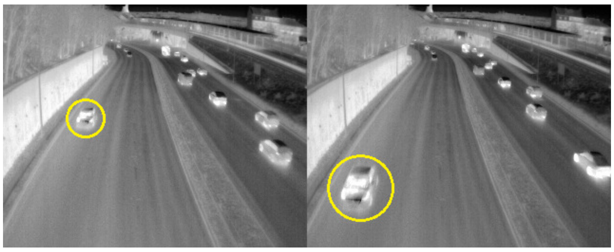


Figure 10: Two pictures, right taken exactly 1 second after left

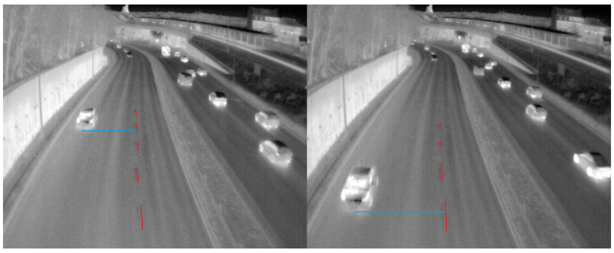


Figure 11: Highlighted stripes (red) and car's front bumper position in regard to the stripes (blue)

V. Future possibilities

Dataset presented in this paper can be used to develop computer vision systems for numerous applications, like traffic monitoring, traffic management for smart cities as well as surveillance and advanced driver assistance systems. Thanks to usage of long-wave infrared imaging, such systems will be immune to even the harshest lighting conditions, providing the same level of accuracy in day and night.

VI. Conclusion

Main contribution of this paper is novel thermal imaging dataset with automotive scenes. It contains several thousands images with hand annotated objects. Second contribution of this paper is trained object detector based on YOLOv5 architecture that shows high accuracy in small object detection alongside with the high speed of operation. Both, the automotive thermal database, as well as the pretrained automotive thermal object detection model, are available free for further research on our website. Moreover, we present sample application of our model for car speed measurement based on thermal images. Clear advantages of such approach are also presented. Finally, in the future we plan to further extend our database, as well as develop more resilient trackers that can reliably operate in dense road conditions and with thermal images.

Acknowledgments

This work was supported by the National Centre for Research and Development, Poland, under the grant no. POIR.01.01.01-00-1116/20.

References

[1] J. M. Lloyd, *Thermal imaging systems*. Springer Science & Business Media, 2013.

[2] M. Krišto, M. Ivasic-Kos, and M. Pobar, "Thermal object detection in difficult weather conditions using yolo," *IEEE access*, vol. 8, pp. 125 459–125 476, 2020.

[3] J. Gąsienica-Józkowy, M. Knapik, and B. Cyganek, "An ensemble deep learning method with optimized weights for drone-based water rescue and surveillance," *Integrated Computer-Aided Engineering*, vol. 28, pp. 221–235, 2021, 3.

[4] M. Knapik and B. Cyganek, "Fast eyes detection in thermal images," *Multimedia Tools and Applications*, vol. 80, no. 3, pp. 3601–3621, 2021.

[5] J. Redmon and A. Farhadi, "Yolo9000: Better, faster, stronger," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 6517–6525, 2017.

[6] ——, "Yolov3: An incremental improvement," *ArXiv*, vol. abs/1804.02767, 2018.

[7] G. Jocher, A. Chaurasia, A. Stoken, J. Borovec, NanoCode012, Y. Kwon, TaoXie, J. Fang, imyhxy, K. Michael, Lorna, A. V, D. Montes, J. Nadar, Laughing, tkianai, yxNONG, P. Skalski, Z. Wang, A. Hogan, C. Fati, L. Mammana, AlexWang1900, D. Patel, D. Yiwei, F. You, J. Hajek, L. Diaconu, and M. T. Minh, "ultralytics/yolov5: v6.1 - TensorRT, TensorFlow Edge TPU and OpenVINO Export and Inference," Feb. 2022. [Online]. Available: https://doi.org/10.5281/zenodo.6222936

[8] M. Bhattarai and M. Martinez-Ramon, "A deep learning framework for detection of targets in thermal images to improve firefighting," *IEEE Access*, vol. 8, pp. 88 308–88 321, 2020.

[9] J. Gong, J. Zhao, F. Li, and H. Zhang, "Vehicle detection in thermal images with an improved yolov3-tiny," in *2020 IEEE international conference on power, intelligent computing and systems (ICPICS)*. IEEE, 2020, pp. 253–256.

[10] W. Zhou, Q. Guo, J. Lei, L. Yu, and J.-N. Hwang, "Ecffnet: Effective and consistent feature fusion network for rgb-t salient object detection," *IEEE Transactions on Circuits and Systems for Video Technology*, 2021.

[11] X. Dai, X. Yuan, and X. Wei, "Tirnet: Object detection in thermal infrared images for autonomous driving," *Applied Intelligence*, vol. 51, no. 3, pp. 1244–1261, 2021.

[12] R. E. Rivadeneira, A. D. Sappa, and B. X. Vintimilla, "Thermal image super-resolution: A novel architecture and dataset." in *VISIGRAPP (4: VISAPP)*, 2020, pp. 111–119.

[13] S. R. Yeduri, D. S. Breland, S. B. Skriubakken, O. J. Pandey, and L. R. Cenkeramaddi, "Low resolution thermal imaging dataset of sign language digits," *Data in Brief*, vol. 41, p. 107977, 2022.

[14] M. Knapik and B. Cyganek, "Driver's fatigue recognition based on yawn detection in thermal images," *Neurocomputing*, vol. 338, pp. 274–292, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231219302280

[15] M. A. Farooq, P. Corcoran, C. Rotariu, and W. Shariff, "Object detection in thermal spectrum for advanced driver-assistance systems (adas)," *IEEE Access*, vol. 9, pp. 156 465–156 481, 2021.

[16] "Darklabel annotation software," https://github.com/darkpgmr/DarkLabel, accessed: 2022-06-07.

[17] "Dz. u. 26.11.2019, position 2311, szczegółowe warunki techniczne dla znaków drogowych poziomych i warunki ich umieszczania na drogach, section 2.2.1.2," https://sip.lex.pl/akty-prawne/dzu-dziennik-ustaw/szczegolowe-warunki-techniczne-dla-znakow-i-sygnalow-drogowych-oraz-17066287, accessed: 2022-06-07.

1^{st} Workshop on Personalization and Recommender Systems

RECOMMENDER Systems are present in our everyday life while we reading news, logging in to social media or buying something at e-shops. Thus, it is not suprising that this domain is getting more and more attention from researchers from academia as well as from industry practitioners. However, the way in which they look at the same problem differs a lot.

Personalization is an important element in novel recommendation techniques. Nonetheless, it is a wider topic that concerns also user modelling and representation, personalized systems, adaptive educational systems or intelligent user interfaces.

The objective of PeRS is to extend the state-of-the-art in Personalization and Recommender Systems by providing a platform at which industry practitioners and academic researchers can meet and learn from each other. We are interested in high quality submissions from both industry and academia on all topics related to Personalization and Recommender Systems.

TOPICS

The list of topics includes, but is not limited to:

- Personalization
  - User Profiles
  - Ontology-based user models
  - Personalized systems
  - Intelligent user interfaces
- Recommender Systems approaches
  - Collaborative Recommender Systems
  - Semantic-based Recommender Systems
  - Context-aware Recommender Systems
  - Cross-domain Recommender Systems
- Machine Learning techniques for Recommender Systems
  - Association Rules
  - Clustering methods
  - Neural Networks
  - Deep Learning
  - Reinforcement Learning
- Applications of Recommender Systems methods
  - News recommendations
  - Tourism recommendations
  - Fashion recommendations
  - Podcasts recommendations
  - Medical recommendations
  - Other domain-specific recommenders
- Evaluation of Recommender Systems
  - Metrics
  - Evaluation studies
  - Reproducibility of existing methods
  - Case studies of real-world implementations

TECHNICAL SESSION CHAIRS

- **Karpus, Aleksandra,** Gdańsk University of Technology, Poland
- **Przybyłek, Adam,** Gdańsk University of Technology, Poland

PROGRAM COMMITTEE

- **Anelli, Vito Walter,** Politechnic University of Bari, Italy
- **Aziz Butt, Shariq,** University of Lahore, Pakistan
- **Borg, Markus,** SICS Swedish ICT AB, Sweden
- **Brzeski, Adam,** Gdańsk University of Technology, Poland
- **Cellary, Wojciech,** WSB Universities in Poznan, Poland
- **Dedabrishvili, Mariam,** International Black Sea University, Georgia
- **de Gemmis, Marco,** University of Bari "Aldo Moro", Italy
- **Dutta, Arpita,** National University of Singapore, Singapore
- **Ghofrani, Javad,** University of Lübeck, Germany
- **Goczyła, Krzysztof,** Gdańsk University of Technology, Poland
- **Inayat, Irum,** National University of Computer and Emerging Sciences, Pakistan
- **Lops, Pasquale,** University of Bari "Aldo Moro", Italy
- **Madeyski, Lech,** Wroclaw University of Technology, Poland
- **Marcinkowski, Bartosz,** University of Gdańsk, Poland
- **Misra, Sanjay,** Ostfold University College, Halden, Norway
- **Mohapatra, Durga Prasad,** NIT Rourkela, India
- **Mukta, Saddam Hossain,** United International University, Bangladesh
- **Ng, Yen Ying,** Nicolaus Copernicus University, Poland
- **Nguyen, Phuong T.,** University of L'Aquila, Italy
- **Nocera, Francesco,** Politechnic University of Bari, Italy
- **di Noia, Tommaso,** Politechnic University of Bari, Italy
- **Orłowski, Cezary,** WSB University in Gdańsk, Poland
- **Polignano, Marco,** University of Bari "Aldo Moro", Italy
- **Poniszewska-Maranda, Aneta,** Lodz University of Technology, Poland

- **Szymański, Julian,** Gdańsk University of Technology, Poland
- **Taweel, Adel,** Birzeit University, Palestine
- **Theobald, Sven,** Fraunhofer IESE, Germany
- **Tkalcic, Marko,** University of Primorska, Slovenia
- **Vagliano, Iacopo,** Amsterdam University Medical Center, Netherlands
- **Wrycza, Stanisław,** University of Gdańsk, Poland

DOI: 10.15439/2022F187
ISSN 2300-5963 ACSIS, Vol. 31

MyMigrationBot: A Cloud-based Facebook Social Chatbot for Migrant Populations

Karol Chlasta, Paweł Sochaczewski, Izabela Grabowska, Agata Jastrzębowska
Kozminski University
CRASH Center for Research on Social Change and Human Mobility
ul. Jagiellońska 57/59, 03-301 Warsaw, Poland
Email: kchlasta@kozminski.edu.pl

***Abstract*—We present the design, implementation and evaluation of a new cloud-based social chatbot called MyMigrationBot, that is deployed to Facebook. The system asks and answers questions related to user's personality traits and person-job competency fit to give feedback, and potentially support migrant populations. The chatbot's response database is based on reputable socio-psychological tools and can be customised. The system's backend is written with Node.js, deployed to AWS and Twilio, and joined with Facebook through Graph and Messenger APIs. To our knowledge this is the first multilingual social chatbot deployed to Facebook and designed to research and support migrant populations with feedback in Europe. It does not have personality like other bots, but it can study and feedback on migrants' personality and on other customised questionnaires e.g., job-competency fit. The aim of a social chatbot in our research project is to help engage migrants with social research using feedback information tailored to them. It can help migrants to get knowledge about their psycho-social resources and therefore to facilitate their integration process into a receiving labour market. We evaluated the chatbot on a group of 53 people, incl. 23 migrants, and we present the results.**

I. Introduction

CHATBOTS are increasingly popular in everyday interactions. The ever increasing affordability and popularity of ICT devices leads to a wider range of communication channels available for different social groups. Is might be viable to use these channels for a number of reasons. Firstly, a social chatbot as an interaction system is able to gather feedback from its respondents and use it to increase the availability of information, the quality of both commercial and public services, and perhaps improve users' quality of life. Secondly, a social chatbot can give personalised feedback to respondents which can increase motivation to participate in research. And thirdly, a social chatbot with feedback can give higher satisfaction out of participating in research [13].

This manuscript focuses on *MyMigrationBot*, a social chatbot system able to gather Big Data for social research from Facebook, and interact with users through multiple communication channels using a Facebook Messenger protocol.

II. Motivation

Migrant populations often lack support, usually from available public services in a receiving country, as there are massive movements in a short period of time, due to some unpredictable circumstances like war or a natural disaster. Such people are usually not prepared to start a new working life in a new labour market.

In recent years Poland, as well as other European countries received millions of both labour immigrants and war refugees. We believe these people could be supported in the receiving societies not only by human agents, but also by new technologies like personal assistants and avatars, often simply known as the social chatbots.

In 2022 a new massive wave of immigration from Ukraine arrived in Poland. As of the 5th of May 2022 according to Operational Data Portal of UN High Commissioner for Refugees[1] 3,143,550 Ukrainians entered Poland out of 5,757,014 people who left Ukraine since February 2022, when the Russian invasion had started, which consists of nearly 55 per cent of all recently fleeting Ukrainian population.

Immigrants and refugees widely use mobile phones and other smart devices. The poly-media accessibility in one smartphone makes it 'a must-to-be taken' by every human on the move [15]. As Dekker et al. (2018) [4] show in relation to Syrian refugees who applied for asylum in the European Union (EU) member states in 2015 and 2016 (the largest group), they used smartphones and social media in migration decision making. The most meaningful issues for refugees as reported by the authors are linked to the access and evaluation of the trustworthiness of information. The kind of social chatbot which we propose in this article could for instance help them to diagnose their psycho-social capitals and to facilitate navigation through the available public services upon arrival.

This is because our chatbot can use different structured or semi-structured questionnaires in its conversations. Apart from that, new custom conversations can be designed to enhance the existing functionality of the *MyMigrationBot*. One of the biggest advantages of our chatbot is the functionality of feedback given to our users as a 'social and non-material' complimentary thank you for their participation. The feedback can be also customised.

To summarise, the main motivation for us was to create a multilingual social chatbot system - as a conversational agent - able to provide migrant populations with targeted, immediate

[1] https://data2.unhcr.org/en/countries/

feedback, so that they can better integrate into a receiving society.

III. Questions and Feedback

The first set of questions asked by *MyMigrationBot* is taken from the Ten-Item Personality Inventory (TIPI) [8] measuring the Big Five personality dimensions: (1) Extroversion, (2) Agreeableness, (3) Conscientiousness, (4) Emotional Stability, and (5) Openness to Experience. As the original TIPI questionnaire was created in English only, we used the Polish adaptation by Sorokowska (2014) [21], and a translation to Ukrainian prepared by a native speaker of Ukrainian. Kosinski at al. (2014) [12] show manifestations of user personality in website choice and behaviour on online social networks, which was a great inspiration for us.

Personality has been linked to migration for a number of reasons. The work offered by social psychologists goes a long way to show that the migrant personality matters in making migration decisions. Personality traits influence international voluntary migration. They also inform about self-selection to migration. Boneva and Frieze (2001) [2] show that individuals who want to emigrate possess a syndrome of personality characteristics that differentiates them from those who want to stay in their country of origin. They also explain the role of personality in desires to emigrate. Emigrants are not just responding to a particular set of economic conditions; there is something specific about the personality of those who desire to move. Emigrants are less prone to anxiety and insecurity than non-emigrants [19]. Higher persistence, openness to experience, as well as previous experience of living internationally, all increased the chances that a participant was planning to move abroad. Higher agreeableness and conscientiousness lowered the odds of a move. Men who were lower in emotional stability were more likely to want to leave, but the same effect was not found for women [22]. Liable et al. (2021) [14] showed that non-cognitive personality traits explain the wage gap between male migrants and non-migrants. Polek et al. (2011) [18] report that personality traits bring us information about adjustments of migrants to the new environments, and they can help in migrant integration processes.

Yet another tool that we incorporated into our *MyMigrationBot* was a Job-Competency Fit Scale [11], which also gives tailor-made feedback to our respondents. This measure of fit is referred to in psychology as a molar measure, a direct question [5]. Although a competence has many meanings, the main meaning is about performing a task, or a function on an adequate level, with knowledge and understanding of a situation. Job fit is the key factor in being hired today in the labour market. We endowed our social chatbot with the list of 26 human competencies connected to cognitive and soft/social competencies which were tested initially in big social surveys on human capital [9]. Then we asked our users: "Think about the job you are currently working in. Then, please specify to what extent the competencies and skills listed below are needed (required) for the position you hold?". Then the chatbot gives feedback according to means verified before. Therefore our social chatbot offers a tool which can help to match a person and their competencies to a proper job.

User experience is a key concept for designing and enhancing the quality and usability of software products [23]. Therefore, we have decided to include System Usability Scale [1] (SUS) to gather some subjective assessment on the *MyMigrationBot*'s design and usability for migrants and non-migrants. It was a simple survey with a standard scale consisting of a 10 item questionnaire with five response options for respondents: from strongly agree to strongly disagree (Likert's social scale). We need to learn more about the bot's accuracy, task completion and responsiveness. User experience evaluation allows us to identify how our MyMigrationBot should evolve in the future to meet experience and expectations [17] of our target population - various categories of migrants. To a certain degree, the assessment feedback by users embeds us into a co-designing process [3], especially since our MyMigrationBot is still in a prototype phase.

Nowadays many chatbots have been designed with an aim to assist with information-seeking, and guidance are based on a frequently asked question and answers mode (FAQ). Sansonett et al. argue that human users expect chatbots to understand their texts, provide adequate answers and that they will be interactive with humans in run-time. As far as we know there are many social chatbots to support psychotherapy (e.g. Woebot) and AI legal services. There are chatbots who are given personality (e.g. XiaoIce of Microsoft) in order to make them user-friendly and responsive to users' queries and interests. There is also a Tinka chatbot - a dialogue system performing as an assistant in the mobile phone delivery system of T-mobile Austria. Tinka is able to act and deliver customer information in various topics. Our benchmark might be Eike, a chatbot (with an avatar) with a personality designed to deliver information to various groups of migrants [3]. According to respondents' testing this chatbot, Eike "should be be a gentle city-born messenger pursuing peace in the neighbourhood. Eike should be able to know about living in a German city and be happy to share their knowledge with anyone who comes to seek it. Eike should soothe worries in a soft and friendly voice and always have a positive appearance, rendering migrants hopeful and optimistic in terms of living and working" (2020: 224). Our approach presented in this article is different. We design, implement and evaluate a social chatbot called *MyMigrationBot* which helps users to learn more about their personality (a bot does not have a personality of its own), their Job-Competency Fit and possible other customised questionnaires with feedback diagnosing psycho-social capital of humans. Therefore our conversation design is slightly different than 'a standard' chatbot design. Next to introduction, greetings, admitting errors, it delivers answers. The bot is particularly designed for migrants to help them to diagnose themselves for the labour market's needs.

IV. Implementation

The architecture of *MyMigrationBot* comprises of front-end in Facebook Messenger, and back-end deployed into AWS

Cloud. The conversation engine uses Twilio [10] platform, that is linked to Facebook using Facebook Send API[2]. All data is stored with AWS RDS service with MySQL 8 engine. The source code of the *MyMigrationBot* is stored in a private GitHub repository[3], shareable with interested parties on-demand. Back-end is written with Node.js version 16.13.1, and it uses Crypto package for Facebook application secrets encryption. We use Node.js MySQL drivers (version 2.18.1) for accessing the database[4], and the Node.js request package[5] (version 2.88.2) for processing https requests to Facebook's Graph API endpoint (in version 12). Node.js Twilio package[6] (version 3.76.1) is used to invoke calls to Twilio Autopilot API endpoint. Both Twilio Autopilot and Twilio Functions were created manually using the Twilio console. We also used Node.js Twilio-cli package[7] to manage our Twilio resources (e.g. Autopilot and Functions).

A. *Dialogue Management*

To drive the responses of *MyMigrationBot* we used Twilio Autopilot, a response and dialogue management system based on customizable, serverless functions from the Twilio cloud platform. The Autopilot service [7] allows us to "build, train, and deploy AI bots using natural language understanding and machine learning". It also simplifies deployment of the required functions from the code repository to different cloud-based environments.

The dialogue manager functionality within Twilio Autopilot chooses which response to return to the user. The system will choose "the best response", so the one that was ranked highest by the engine. If the score of the top ranked response is below a defined threshold (determined and customizable in the Autopilot's configuration), the dialogue manager system will select an off-topic response instead, that indicates lack of understanding (e.g. say "I do not understand, please repeat."). The system also contains a simulator allowing to both test and train the chatbot.

Apart from Twilio, the *MyMigrationBot* conversations can be monitored through the administration panel of a Facebook Page, to which it is linked. Moreover, a Facebook user granted page administrator role can monitor and participate in chatbot's conversations. As a result a human agent is able help users facings issues in conversations, to help them progress through the survey, or to help the chatbot in responding to any non-standard interactions, e.g. to avoid biased contents of databases [20].

While Twilio platform has recently been used for custom chat applications using WhatsUp by Immigration Policy Lab at Stanford University [6], this is the first deployment of Twilio cloud platform with Facebook Messenger to migration studies, and second to support migrants in Europe.

B. *Facebook APIs*

Facebook launched the Messenger platform in 2016[8]. We use Facebook Messenger API through Twilio platform. The API connects to the backend of *MyMigrationBot*, which was deployed to AWS Cloud, and it bridges Twilio with Facebook. When a message event occurs, it notifies bot's web-hook and calls a predefined Twilio function.

We also use Facebook Graph API[9] in its latest version (v13.0). As stated by Facebook, the Graph API is the primary way to read and write to the Facebook social graph, and all their SDKs and products "interact with the Graph API in some way". We used this API to retrieve a limited set of demographic information about *MyMigrationBot*'s users from their public Facebook profiles. We recognise the fact that due to recent controversies and data leaks [16], the use of Facebook's Graph API is now made more difficult for any non-public elements of the profile. The process now requires a multi-level approval, and a time-consuming vetting process from Meta, Facebook's parent company. This might impact the timelines of adding any new attributes to our database in future.

At the moment, each time a conversation with *MyMigrationBot* is started, the basic set of attributes is saved into the RDS in AWS Public Cloud. These attributes are described in Table II of the Appendix section.

Once all the back-end actions of *MyMigrationBot* are complete, the relevant text message is selected using the dialogue management system, and the response is sent to Facebook to deliver to the users' Messenger client(s). The process ends with matching the user's responses to the variables for the relevant user session. A screenshot of multilingual interaction with *MyMigrationBot* is presented in Figure 1.

C. *Infrastructure*

The infrastructure of this project is hosted in AWS Cloud. A few of the AWS services and their components were created, and configured manually (e.g. Virtual Private Cloud, Public and Private Sub-nets, Network ACLs and IAM users). Other AWS services, like RDS, EC2, Security Groups, Launch configuration, and Auto Scaling groups were deployed automatically using Terraform (version 1.0.11). The Infrastructure was loaded from the code by Terraform using development workstations. Terraform state files are kept in Amazon S3. Terraform gets access to AWS Cloud via AWS IAM user access keys. The same method of authentication is used by AWS Command Line Interface to manage AWS resources for the project.

Our infrastructure in AWS is kept in Virtual Private Cloud (VPC) in Europe (Ireland) region. We used Internet Gateway Component to open network traffic between the public Internet and our VPC.

We protect our infrastructure with two layers of firewall, on the subnet and application level. We use Network ACLs

[2] https://developers.facebook.com/docs/messenger-platform

[3] https://github.com/KarolChlasta/BigMig

[4] https://www.npmjs.com/package/mysql

[5] https://www.npmjs.com/package/request

[6] https://www.npmjs.com/package/twilio

[7] https://twil.io/cli

[8] https://about.fb.com/news/2016/04/messenger-platform-at-f8/

[9] https://developers.facebook.com/docs/graph-api/

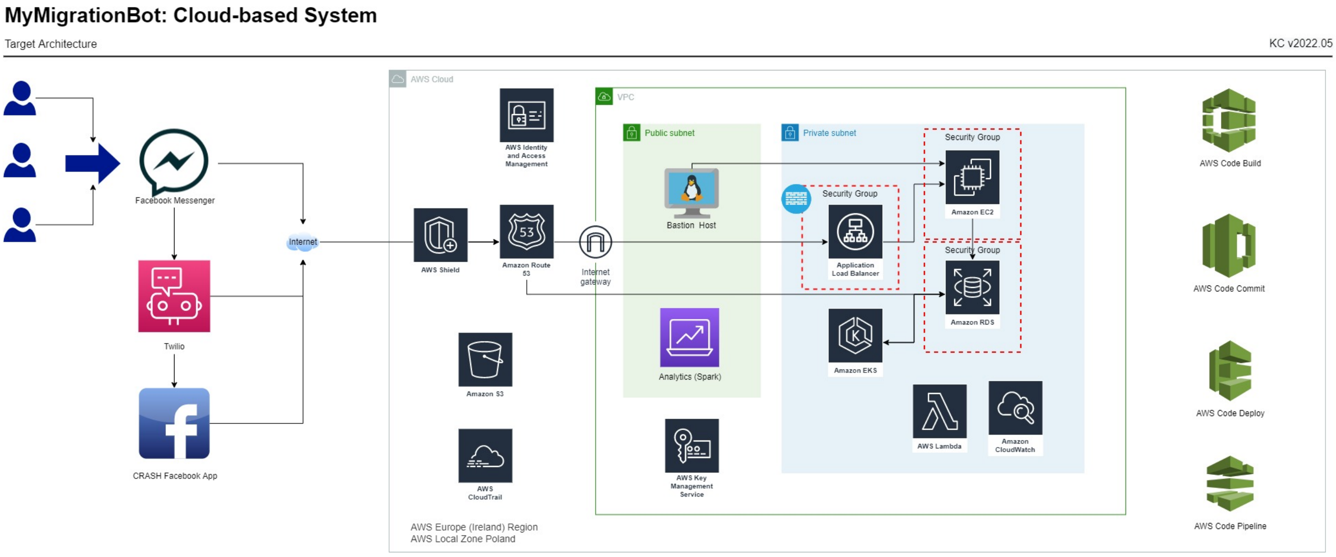


Fig. 1. Target Architecture of MyMigrationBot

for subnet level firewall and Application Firewall for Security Groups, for which we opened the traffic only for the protocols and ports that are used in the project. The Back0end was installed on a single EC2 instance, which was created via Auto Scaling Groups. The Server Instance Type of t2.micro (1 CPU, 1 GiB Memory, 3.3 GHz Intel Scalable Processor) was selected for the beta testing.

To protect our back-end server's webhooks against execution from unauthorised third-party applications we use Facebook Page and Verify Tokens, as well as application secrets.

Apart from having a specific domain name hosted in AWS Route 53, we also used Ngrok service to publish URL for our back-end server. We haven't used the AWS Application Load Balancer, as it was not justified by the beta testing stage of the project.

The back-end server is run as a Windows service. It is started automatically during the boot process of the EC2 instance.

V. Evaluation

A. Participants

We recruited N=53 participants for the study. Mostly from Poland (n = 34; 64.2%), men (n = 33; 62.26), filling the tool on the computer (n = 25; 47.2%) or on the phone (n = 20; 37.7%). Almost half of them are people with experience of migration (n = 23; 43.4%). Mean age (M = 31.02; SD = 12.73). Medium ICT skills (M = 4.15 / 5.00; SD = 0.91). Note that the male group includes a single person, who declares a gender that is not aligned with the person's legal status in the country. Although every participant was an active Facebook user, only 21 (39.6%) declared their ICT skills as excellent. Detailed data on participants is presented in Table I.

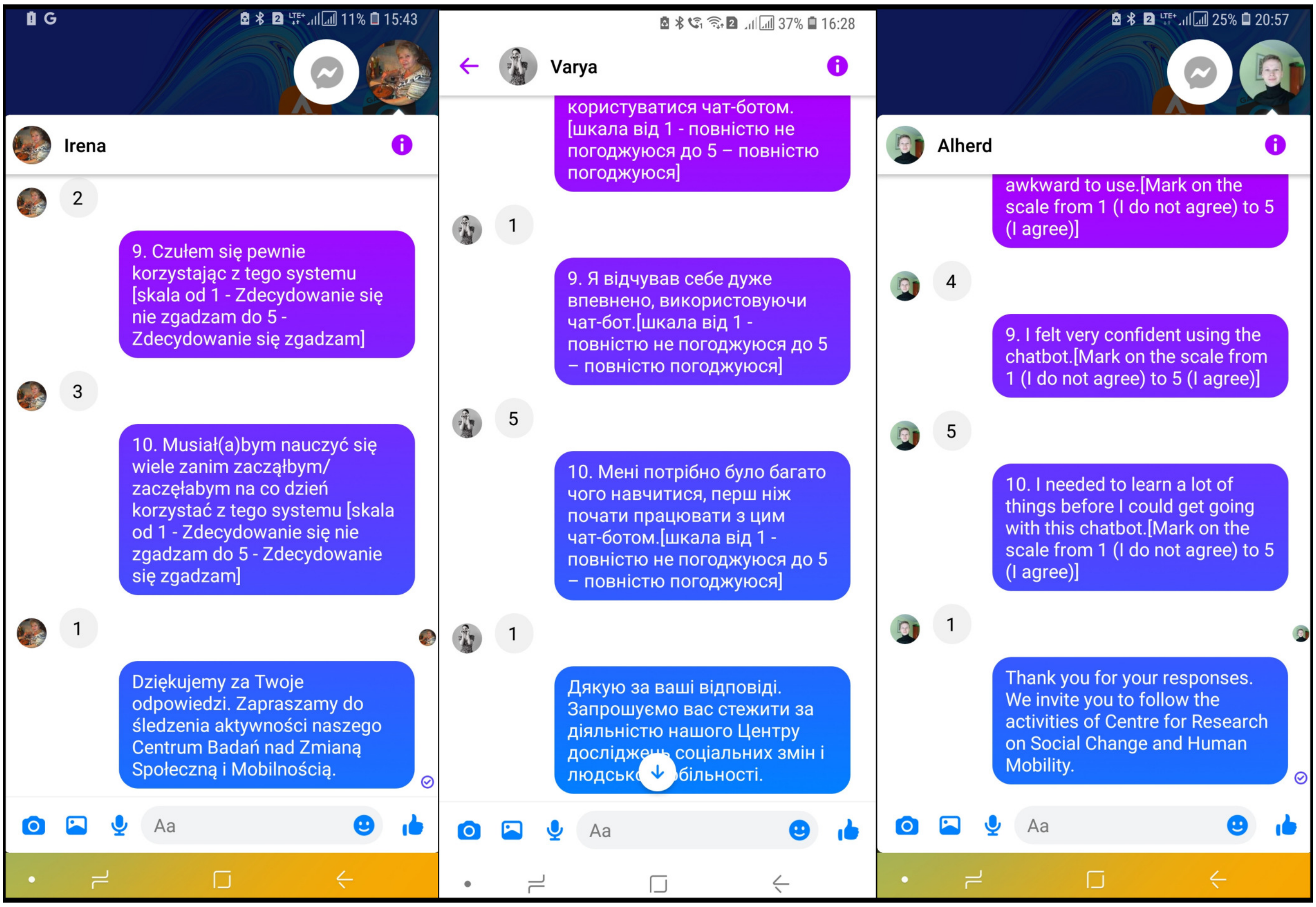
TABLE I
Descriptive Characteristics of the Participants

Variable Name	Category	Number	%
Nationality	Polish	34	64.2
	Ukrainian	6	11.3
	Belarusian	5	9.4
	British	1	1.9
	Czech	1	1.9
	French	1	1.9
	Polish-Italian	1	1.9
	No data	4	7.5
Country of Birth	Polish	34	64.2
	Ukraine	6	11.3
	Belarus	5	9.4
	UK	2	3.8
	Czechia	1	1.9
	France	1	1.9
	India	1	1.9
	No data	3	5.7
Gender	Female	20	37.73
	Male	32	62.26
Device	Computer	25	47.2
	Mobile Phone	21	39.6
	No data	7	13.2
Immigrant	No	30	56.6
	Yes	23	43.4

B. Procedure

The participants were engaged with a pilot study of MyMigrationBot via a Facebook page of CRASH Center for Research on Social Change and Human Mobility of Kozminski University using Facebook Messenger[10].

[10] https://www.facebook.com/CRASHKozUni



Irena

2

9. Czułem się pewnie korzystając z tego systemu [skala od 1 - Zdecydowanie się nie zgadzam do 5 - Zdecydowanie się zgadzam]

3

10. Musiał(a)bym nauczyć się wiele zanim zacząłbym/ zaczęłabym na co dzień korzystać z tego systemu [skala od 1 - Zdecydowanie się nie zgadzam do 5 - Zdecydowanie się zgadzam]

1

Dziękujemy za Twoje odpowiedzi. Zapraszamy do śledzenia aktywności naszego Centrum Badań nad Zmianą Społeczną i Mobilnością.

Varya

користуватися чат-ботом. [шкала від 1 - повністю не погоджуюся до 5 – повністю погоджуюся]

1

9. Я відчував себе дуже впевнено, використовуючи чат-бот.[шкала від 1 - повністю не погоджуюся до 5 – повністю погоджуюся]

5

10. Мені потрібно було багато чого навчитися, перш ніж почати працювати з цим чат-ботом.[шкала від 1 - повністю не погоджуюся до 5 – повністю погоджуюся]

1

Дякую за ваші відповіді. Запрошуємо вас стежити за діяльністю нашого Центру досліджень соціальних змін і людсько[illegible]більності.

Alherd

awkward to use.[Mark on the scale from 1 (I do not agree) to 5 (I agree)]

4

9. I felt very confident using the chatbot.[Mark on the scale from 1 (I do not agree) to 5 (I agree)]

5

10. I needed to learn a lot of things before I could get going with this chatbot.[Mark on the scale from 1 (I do not agree) to 5 (I agree)]

1

Thank you for your responses. We invite you to follow the activities of Centre for Research on Social Change and Human Mobility.

Fig. 2. *MyMigrationBot*'s interaction with users of mobile devices on Facebook Messenger in Polish, Ukrainian and English

Participants used their own personal computers or mobile devices to interact with *MyMigrationBot*. After completing a single interaction with the chatbot, they were asked to answer a post-study questionnaire on the usability of that interaction, with optional open-ended questions on *MyMigrationBot*'s advantages and disadvantages. In took around 5-10 minutes to complete the study.

VI. Results

System Usability Scale (SUS) was used. This is a 10 item questionnaire with five response options for respondents; from Strongly agree to Strongly disagree. SUS allows us to evaluate a wide variety of products and services, including hardware, software, mobile devices, websites and applications [1]. Though the scores are 0-100, SUS score above 68 would be considered above average and anything below 68 is below average.

Both on the computer and on the phone, the *MyMigrationBot* respondents' results at SUS were in both cases about $M = 70.00$ (computer $M = 70.00$ and mobile phone $M = 70.25$).

To compare the System Usability Scale between respondents who experienced and did not experience migration, t-Student for independent groups analysis was counted. It turned out that people with no experience of migration assessed the tool usability slightly better ($M = 71.08; SD = 8.14$) than people with experience of migration ($M = 68.26; SD = 12.14$). However, this difference turned out to be statistically insignificant $t(51) = 1.012; p > 0.05$.

We have also gathered additional, optional, unstructured feedback from the participants in Table III of the Appendix section of this manuscript. The feedback was based on 5 open-ended questions.

VII. Discussion

With the rapid inflow of migrants to Europe the social chatbots have their momentum [3] both for research, and for practical use associated with information-seeking and migrant integration as a result.

Our *MyMigrationBot* is not designed as a persona with a personality but as a professional trustworthy assistant to diagnose a respondent's personality and other psycho-social resources through an individual tailor-made feedback protocol. The architecture of our bot is flexible enough to customise questionnaires with feedback when new needs emerge. The main limitation is that we have not been able to fully implement the target architecture in relation to all the AWS cloud components and automation services. As a result, the maintenance of the *MyMigrationBot* is still partially manual. The configuration of the server infrastructure (AWS EC2, Ngrok) was automated using the user data script that gets executed on the instance's initial boot. The script downloads and installs all the required utilities e.g. AWS Client, Ngrok, as well as retrieving the source code from GitHub, but that script is still triggered from the developers' laptops.

At present we also do not use all the services listed on Figure 1, namely the Analytics component (Spark), as well as the AWS CodePiplines, CodeBuild, and CodeDeploy. We use a free GitHub repository instead. The reason for that limitation was cost. Our intention was to cut the cost of the project until the beta testing finishes. All the cloud infrastructure is attached to the main author's individual Twilio and AWS accounts (and his credit card). This was also the reason for a technical decision to host our Internet backend with Ngrok service.

The *MyMigrationBot* was also tested on limited number of users, with the maximum of 5 users using the system at the same time on April 28th 2022. We recognise that more in depth performance testing is needed prior to the production release of the application.

Additionally, manual configuration of Twilio Autopilot and Twilio Functions is prone to human error. In future, we want to be able to automate creation of Autopilot (together with Twilio Services, that replace Twilio Functions) directly from the source code, using Twilio CLI. We believe this will simplify building new chatbots, and it will be convenient for multi-chatbot environments. We have already implemented this approach, but have not been able to fully test it. Another social limitation concerns the fact that the bot itself does not recruit survey respondents, which is currently the biggest challenge for social and marketing research.

VIII. Conclusion

MyMigrationBot is the first deployment of a Facebook social chatbot using Twilio in Central Europe, and the second in Europe [3] to support migrant population. To our knowledge, it is also the first chatbot using Facebook's Graph API for information retrieval, to gather reliable Big Data for social science research, with a special focus on migrant populations.

The application was positively assessed by our beta testers, who were both migrants and stayers. There was no significant statistical difference in perceived usability between the groups. People do not like filling out surveys but can be motivated by constructive feedback given to them [13]. We have gathered a lot of interesting feedback from 53 beta testers. We will focus our efforts on improving user experience e.g. by adding graphical elements to the surveys, and overcoming the key limitations highlighted by our users in beta testing, prior to publicly releasing the *MyMigrationBot* for migrant populations. We also plan to place our MyMigrationBot into a co-design process with migrants in focused groups [3] in order to: firstly, maximise an interactive engagement of our users and therefore a conversation flow; secondly, to enhance the usability of the machine; and thirdly, to refine our initially proposed solution. We also plan to deliver our *MyMigrationBot* to migrant population as part of a platform for assessments of various psycho-social capitals of migrants which are used in the labour market.

Acknowledgment

We thank Jennifer Fei and Jessica Sadye Wolff from Immigration Policy Lab at Stanford University and Michał Kosiński from Stanford Graduate Business School for several discussions that led to the development of *MyMigrationBot*. We thank Michael Connolly for proofreading, Dominika Winogrodzka and Ivanna Kyliushyk for their ongoing support with *MyMigrationBot*. The publication was supported by a grant from National Science Centre (Poland) no. 2020/37/B/HS6/02342.

References

[1] Aaron Bangor, Philip T Kortum, and James T Miller. "An empirical evaluation of the system usability scale". In: *Intl. Journal of Human–Computer Interaction* 24.6 (2008), pp. 574–594.

[2] Bonka S Boneva and Irene Hanson Frieze. "Toward a concept of a migrant personality". In: *Journal of Social Issues* 57.3 (2001), pp. 477–491.

[3] Zhifa Chen et al. "Creating a chatbot for and with migrants: chatbot personality drives co-design activities". In: *Proceedings of the 2020 ACM Designing Interactive Systems Conference*. 2020, pp. 219–230.

[4] Rianne Dekker et al. "Smart refugees: How Syrian asylum migrants use social media information in migration decision-making". In: *Social Media+ Society* 4.1 (2018), p. 2056305118764439.

[5] Jeffrey R Edwards et al. "The phenomenology of fit: linking the person and environment to the subjective experience of person-environment fit." In: *Journal of applied psychology* 91.4 (2006), p. 802.

[6] Jennifer Fei et al. "Automated Chat Application Surveys Using WhatsApp". In: (2020).

[7] Ashley Firth. "New Technologies". In: *Practical Web Inclusion and Accessibility*. Springer, 2019, pp. 355–392.

[8] Samuel D Gosling, Peter J Rentfrow, and William B Swann Jr. "A very brief measure of the Big-Five personality domains". In: *Journal of Research in personality* 37.6 (2003), pp. 504–528.

[9] Izabela Grabowska and Agata Jastrzebowska. *Migration and the Transfer of Informal Human Capital: Insights from Central Europe and Mexico*. Routledge, 2022.

[10] Srini Janarthanam. *Hands-on chatbots and conversational UI development: build chatbots and voice user interfaces with Chatfuel, Dialogflow, Microsoft Bot Framework, Twilio, and Alexa Skills*. Packt Publishing Ltd, 2017.

[11] Agata Jastrzębowska et al. *Dopasowanie kompetencyjne człowieka do pracy*. Wydawnictwo Naukowe Scholar Sp. z o.o., 2020.

[12] Michal Kosinski et al. "Manifestations of user personality in website choice and behaviour on online social networks". In: *Machine learning* 95.3 (2014), pp. 357–380.

[13] Simon Kühne and Martin Kroh. "Personalized feedback in web surveys: Does it affect respondents' motivation and data quality?" In: *Social Science Computer Review* 36.6 (2018), pp. 744–755.

[14] Marie-Christine Laible and Hanna Brenzel. "Does Personality Matter? Noncognitive Skills and the Male Migrant Wage Gap in Germany". In: *International Migration Review* (2021), p. 01979183211037315.

[15] Mirca Madianou. "Smartphones as polymedia". In: *Journal of Computer-Mediated Communication* 19.3 (2014), pp. 667–680.

[16] Ryan M McManaman. "Strategic Audit of Facebook Through the Lens of International Reputation". In: (2019).

[17] Ana M Moreno et al. "HCI practices for building usable software". In: *Computer* 46.4 (2013), pp. 100–102.

[18] Elzbieta Polek, Jan Pieter Van Oudenhoven, and Jos MF Ten Berge. "Evidence for a "migrant personality": Attachment styles of Poles in Poland and Polish immigrants in the Netherlands". In: *Journal of Immigrant & Refugee Studies* 9.4 (2011), pp. 311–326.

[19] John J Ray. "The traits of immigrants: A case study of the Sydney Parsees". In: *Journal of Comparative Family Studies* 17.1 (1986), pp. 127–130.

[20] Ari Schlesinger, Kenton P O'Hara, and Alex S Taylor. "Let's talk about race: Identity, chatbots, and AI". In: *Proceedings of the 2018 chi conference on human factors in computing systems*. 2018, pp. 1–14.

[21] Agnieszka Sorokowska et al. "Polska adaptacja testu Ten Item Personality Inventory (TIPI)–TIPI-PL–wersja standardowa i internetowa". In: (2014).

[22] Aidan S Tabor, Taciano L Milfont, and Colleen Ward. "The migrant personality revisited: Individual differences and international mobility intentions". In: *New Zealand Journal of Psychology (Online)* 44.2 (2015), p. 89.

[23] Arnold POS Vermeeren et al. "User experience evaluation methods: current state and development needs". In: *Proceedings of the 6th Nordic conference on human-computer interaction: Extending boundaries*. 2010, pp. 521–530.

APPENDIX

TABLE II
MYMIGRATIONBOT'S LOGICAL DATA MODEL SUMMARY

Attribute Name	Type	Default	Source	Description
Id (PRIMARY KEY)	bigint	NOT NULL	Database	Primary key of the record in database
Fb_Id	varchar(255)	NOT NULL	Messenger protocol	Facebook user identifier
First_name	varchar(255)	NULL	Graph API	FB user's first name
Last_name	varchar(255)	NULL	Graph API	FB user's last name
Locale	varchar(255)	NULL	Graph API	FB user's home address
Hometown	varchar(255)	NULL	Graph API	FB user's birth place
Timezone	float	NULL	Graph API	FB user's current timezone offset from UTC
Birthday	varchar(255)	NULL	Graph API	FB user's birthday
Gender	varchar(255)	NULL	Graph API	FB user's gender
TIPIPL_odp_1	smallint	NULL	User input	User answer for TIPIPL question
TIPIPL_odp_2	smallint	NULL	User input	User's answer for TIPIPL question
TIPIPL_odp_3	smallint	NULL	User input	User's answer for TIPIPL question
TIPIPL_odp_4	smallint	NULL	User input	User's answer for TIPIPL question
TIPIPL_odp_5	smallint	NULL	User input	User's answer for TIPIPL question
TIPIPL_odp_6	smallint	NULL	User input	User's answer for TIPIPL question
TIPIPL_odp_7	smallint	NULL	User input	User's answer for TIPIPL question
TIPIPL_odp_8	smallint	NULL	User input	User's answer for TIPIPL question
TIPIPL_odp_9	smallint	NULL	User input	User's answer for TIPIPL question
TIPIPL_odp_10	smallint	NULL	User input	User's answer for TIPIPL question
TIPIPL_ekstarwersja	double(5,1)	NULL	Backend	Evaluation of user's answer for Extraversion
TIPIPL_ugodowosc	double(5,1)	NULL	Backend	Evaluation of user's answer for Agreeableness
TIPIPL_sumiennosc	double(5,1)	NULL	Backend	Evaluation of user's answer for Conscientiousness
TIPIPL_stabilnosc	double(5,1)	NULL	Backend	Evaluation of user's answer for Emotional Stability
TIPIPL_otwartosc	double(5,1)	NULL	Backend	Evaluation of user's answer for Openness to Experience
DopKomp_czy_pracujesz	varchar(5)	NULL	User input	Are you employed? (Flag)
DopKomp_odp_num_1	varchar(5)	NULL	User input	User's answer to Competence-Job fit question
Inter_odp_1	smallint	NULL	User input	User's answer to System Usability Scale
Inter_odp_2	smallint	NULL	User input	User's answer to System Usability Scale
Inter_odp_2	smallint	NULL	User input	User's answer to System Usability Scale
Inter_odp_3	smallint	NULL	User input	User's answer to System Usability Scale
Inter_odp_4	smallint	NULL	User input	User's answer to System Usability Scale
Inter_odp_5	smallint	NULL	User input	User's answer to System Usability Scale
Inter_odp_6	smallint	NULL	User input	User's answer to System Usability Scale
Inter_odp_7	smallint	NULL	User input	User's answer to System Usability Scale
Inter_odp_8	smallint	NULL	User input	User's answer to System Usability Scale
Inter_odp_9	smallint	NULL	User input	User's answer to System Usability Scale
Inter_odp_10	smallint	NULL	User input	User's answer to System Usability Scale
Record_created	datetime	NULL	Backend	Database date and time of record creation
Jezyk	varchar(10)	NULL	User input	Language flag for the surveys executed
Profile_pic	varchar(500)	NULL	Graph API	User's Facebook profile picture (Url)
Age	smallint	NULL	Graph API	Age of Facebook user
It_skils	smallint	NULL	Graph API	FB user's level of it skils (1-5)
Immigrant	smallint	NULL	Graph API	FB user's immigrant flag
Device	varchar(500)	NULL	Graph API	FB user's device type flag

TABLE III
PARTICIPANTS' FEEDBACK ON MYMIGRATIONBOT

What are your overall impressions of using this chatbot on Facebook?
- "Generally a good impression. Easy to use, readable."
- "Chatbot was ok to use, the questions were easy to understand. There were no technical problems."
- "It was fun, the questions were interesting. It gave me something to think about."
- "Easy to use, horrible in visual design."
- "Pretty good, sometimes the languages were not consistent"
- "Quite short, and not a bad experience overall."
- "When I inputted a random letter, the chatbot corrected me – asked to use one of 5 valid answers, which is nice to see."
- "The program is friendly and convenient, the survey can be completed anywhere, and only takes a short while."
- "At first I did not know how to start, nothing was displayed until I made the first move. I had to initiate the conversation."
- "I didn't know what to do, when the greeting in three languages came out."
- "It gives a person a feeling of communication with a person, and therefore some comfort."
What are the advantages of the chatbot experience/interaction?
- "Easy to understand questions, variety of languages."
- "It was fun to look into ourselves and to see how we are."
- "The questions were interesting and it gave me a bit to think about myself and how I behave around people."
- "Not having to scroll through different pages and check boxes, using natural language, even if it was just typing in numbers, was a lot more natural."
- "Easy to use. Fast. Easy to use/read, few languages.It's easily accessible."
- "It was fast, very simple to use, pretty clear and non-sophisticated."
- "Also the feedback (input only values 1-5) in case of a wrong answer was a nice thing."
- "It was nice to see in what category I belong to. It helps us have a better idea of who we are. It wasn't too long to take."
What are the disadvantages of the chatbot experience/interaction?
- "Different scale for each question (sometimes 1-7 or 1-5,1-6) which is confusing and easy to miss, also I do not find this bot helpful, it is not complex enough."
- "When I chose English language, there are some sentences in Polish."
- "Scales across all surveys should be the same. At the moment the psychological tools use the scale from 1-5, 1-6 or from 1 to 7."
- "Executing survey in English, I found a few chatbot answer in Polish."
- "I didn't know how to start the conversation with the social chatbot."
- "There is No ability to change answers."
- "Lack of the option of choosing the answer by clicking the mouse, currently they must type the answer from the keyboard."
- "It's a little more difficult to read through the longer messages."
- "The bot should print the messages in order. Sometimes it sent multiple different messages in the same time (results + questions)."
- "The worst was the fact that in the end there were questions (are you employed) which were put in between other results statements which was very confusing."
- "User should not be given questions and statements in random order."
- "Regrading adding answers, using keyboard is less convenient that clicking on the links in some kind of form."
- "After selecting English as a language – some of the messages were still written in Polish."
- "There was also no clear instruction in the chatbot itself to begin the conversation."
What made it difficult for you to use/interact with the chatbot?
- "Sometimes it was hard put a number on any expression. Other than that I had no difficulties."
- "Really easy to use if you have eyes to read. It's hard to read long messages on the messenger."
- "There were some hiccups with the language the bot used. Even though I selected English at the beginning, parts of the messages were in Polish it was confusing."
- "Nothing, I was only afraid of typing something wrong – I wouldn't like to break the chatbot."
- "Finding questions in such long pieces of text was a bit challenging."
- "In the end the question was combined with a "Thank you for completing. . . .etc." and I didn't see there was a question after all."
- "Lack of instructions Inconsistency in language Messages appearing too quickly.I didn't know if I should read or answer the next ones."
- "The scale (1-7;1-5) was very complicated. In chatbots without visuals maybe it would be easier to have simpler scale like 1-3."
How can we improve the experience/interaction with the chatbot?
- "I would also recommend to add a 'sleep time' for the bot (wait few seconds before the bot sends its response)."
- "In the questions maybe disclose different behaviours, but similar, so that a person while using the chatbot would not feel that they have to choose only one."
- "Instead create more questions that might give you more insight to the person."
- "I would recommend to add "sleep time" for the bot (wait few seconds before boot sends its respond)."
- "Any sign that bot is working, gives a person feeling of communication with a person and therefrom – comfort."
- "Split questions for sure. Some different traits were put together and while I felt 3 with one, I felt 7 with other. So I chose more or less 5 but that wasn't good."
- "In the final part no matter what language you choose you get a sum-up in polish which may be off-putting for some people because they wouldn't understand it."
- "Maybe also split the long blocks of text into smaller ones – especially if there are questions mixed with statements."
- "Instead of having the user type in numbers, use chatbot buttons."
- "Maybe at the end, the chatbot can compile a small file of all the results to send to the user."
- "shorter messages. It could be anonymous."
- "Make buttons to click on, instead of having numbers to type. Reduce the repetitive questions."
- "Unify the language after choosing one."
- "Add intro like: Hi I am a chatbot that will help you to."

Network Systems and Applications

MODERN network systems encompass a wide range of solutions and technologies, including wireless and wired networks, network systems, services, and applications. This results in numerous active research areas oriented towards various technical, scientific and social aspects of network systems and applications. The primary objective of Network Systems and Applications conference track is to group network-related technical sessions and promote synergy between different fields of network-related research.

The rapid development of computer networks including wired and wireless networks observed today is very evolving, dynamic, and multidimensional. On the one hand, network technologies are used in virtually several areas that make human life easier and more comfortable. On the other hand, the rapid need for network deployment brings new challenges in network management and network design, which are reflected in hardware, software, services, and security-related problems. Every day, a new solution in the field of technology and applications of computer networks is released. The NSA track is devoted to emphasizing up-to-date topics in networking systems and technologies by covering problems and challenges related to the intensive multidimensional network developments. This track covers not only the technological side but also the societal and social impacts of network developments. The track is inclusive and spans a wide spectrum of networking-related topics.

The NSA track is a great place to exchange ideas, conduct discussions, introduce new ideas and integrate scientists, practitioners, and scientific communities working in networking research themes.

TOPICS

- Networks architecture
- Networks management
- Quality-of-Service enhancement
- Performance modeling and analysis
- Fault-tolerant challenges and solutions
- 5G developments and applications
- Traffic identification and classification
- Switching and routing technologies
- Protocols design and implementation
- Wireless sensor networks
- Future Internet architectures
- Networked operating systems
- Industrial networks deployment
- Software-defined networks
- Self-organizing and self-healing networks
- Mulimedia in Computer Networks
- Communication quality and reliability
- Emerging aspects of networking systems

Track 3 includes technical sessions:

- Complex Networks—Theory and Application (1$^{\text{st}}$ Workshop CN-TA'22)
- Internet of Things—Enablers, Challenges and Applications (6th Workshop IoT-ECAW'22)
- Cyber Security, Privacy and Trust (3rd International Forum NEMESIS'22)

1^{st} Workshop on Complex Networks: Theory and Application

IN the nature and the world around us, we can observe many network structures that interconnect various elements such as cells, people, urban centers, network devices, companies, manufacturing machines, etc. Most of them have the nature of evolving networks whose structure changes over time. The analysis of such systems from the complex networks point of view allows for better understanding of the processes within them, which can be used to optimize their structure, improve their management methods, detect failures, improve their operating efficiency and plan their development and evolution.

The main goal of this event is to exchange knowledge and experience between specialists from different areas who in their research and design work use theories and solutions characteristic for complex systems. We believe that the meeting will create new ideas and concepts that will affect the development of contemporary methods of design, operation and analysis of network systems.

TOPICS

The list of topics includes, but is not limited to:

- Complex networks architecture
- Large scale networks analytics
- Mathematical and numerical analysis of networks
- Modeling of computer networks
- Cognitive networks
- Visualizations of network processes
- Dynamics on networks
- Biological and physical models on networks
- Dynamic modification of communication protocols parameters for enterprise and ISP systems
- Complex network management
- Performance modeling and analysis in complex networks
- Network function virtualization
- Social networks
- Graph theory and network algorithm application
- Evolving networks
- Detection of anomalies in the functioning of an enterprise-class computer network element
- Predictive maintenance
- Network technologies supporting society 5.0 and education 5.0
- Architecture for next-generation network applications
- Distributed complex systems for remote working and collaboration
- Algorithms for controlling and monitoring complex computer networks

TECHNICAL SESSION CHAIRS

- **Bolanowski, Marek,** Rzeszow University of Technology, Poland
- **Paszkiewicz, Andrzej,** Rzeszow University of Technology, Poland

PROGRAM COMMITTEE

- **Al-Naday, Mays,** University of Essex, United Kingdom
- **Ballas, Rüdiger G.,** Mobile University of Technology, Germany
- **Houssein, Essam H.,** Minia University, Egypt
- **Ignaciuk, Przemysław,** Lodz University of Technology, Poland
- **Kryvyi, Serhii,** Taras Shevchenko National University of Kyiv, Ukraine
- **Kuchanskyy, Vladislav,** Institute of Electrodynamics of the National Academy of Sciences of Ukraine
- **Palau, Carlos,** Universitat Politècnica de València, Spain
- **Provotar, Oleksandr,** Taras Shevchenko National University of Kyiv, Ukraine

Position Papers of the of the 17th Conference on Computer Science and Intelligence Systems pp. 65–68
DOI: 10.15439/2022F280
ISSN 2300-5963 ACSIS, Vol. 31

Evaluation without Ground Truth: a Comparative Study on Preference Mining Techniques in Twitter Social Network

Fabiola S. F. Pereira
Faculty of Computing
Federal University of Uberlândia
Uberlândia, Minas Gerais, Brazil
Email: fabiola.pereira@ufu.br

***Abstract*—In social media research the lack of ground truth for evaluation is a recurrent problem. We study the preference mining task in Twitter network which suffers from this lack of ground truth problem. We implement three different methods from literature, considering a common preference domain of news and carry a comparative study among them. Our preliminary findings show that is possible to combine methods in order to avoid unfeasible user surveying baselines and enable the evaluation of techniques. In the future, our target is to completely eliminate ground truth sets and evaluate based on correlation and causality techniques.**

I. Introduction

In SOCIAL media research evaluation without ground truth is a pressing need [1]. Social networks are characterized by a huge volume of unstructured data, which can reveal from spatiotemporal and causal patterns to users sentiments. The problem is that mining such patterns is challenging due to the lack of reference values previously established. For example, according to [1] researchers are interested in discover when and where certain user activity is likely to occur – when the user is going to search for restaurant reviews? Where the user will be in the evening? Without surveying this user, however, the gap between prediction algorithms and reality can be deep.

We investigate the user preference mining problem in social networks [2]. Specifically, we look for patterns that describe preferences of a given social network user. For instance, considering a domain of news, we want to discovery what are the user u preferred themes through her interactions in her social network. Thus, we can discover that u prefers to read news about *politics* over *sports* news, for example. This task fits exactly in the above discussed problem where we do not have ground truth preference values neither is feasible to manually survey each user about her preferences in the whole social network.

To tackle this problem, we conduct a comparative study among three different preference mining methods in Twitter dataset. The goal is to analyze the behavior of independent models when mining preferences and then overlap results. This resultant intersection set of mined preferences can perform as ground truth. Our contributions are two-fold. (i) Bring a set of different preference mining methods to the same context of Twitter and (ii) compare and overlap results building a trustworthy set of preferences to fill the gap caused by the lack of a ground truth.

II. The User Preference Mining Problem

User preference is a specific type of opinion that establishes an order relation between two objects. For example, when a user says: “I prefer sports than economy”, we clearly identify her preference to sports themes over economy ones. These preference order relations (or preferences, for short) respect the irreflexive and transitive properties. We denote by $o_1 \succ_u o_2$ the preference of user u by the object o_1 over o_2 for $o_1, o_2 \in D$, where D is a preference domain.

Methods for learning and predicting preferences in an automatic way are among the recent research topics in disciplines such as machine learning, knowledge discovery and recommender systems. Approaches relevant to the area range from approximating of an as effective as possible question-answer process (preference elicitation) to collaborative filtering where customer preferences are estimated from the preferences of other customers [2]. In fact, problems of preference learning can be formalized within various settings, depending on the underlying type of preference model or the type of information provided as an input to the learning system. We explore the role of user preferences in recommender systems.

In a general way, to build an effective preference prediction system the following process is executed (Figure 1): first elicit patterns from feedback, which can be explicit (e.g. rating movies) or implicit (e.g. social data, visual perception, clicks, logs). The preference mining task consists in deriving a model from feedback able to infer a preference order between two given objects. This model is often referred to as prediction model. In some proposals, any preference mining task is used, and there is just a user profiling module that seeks to represent preferences through feature vectors or tensors. In the end, given some items and a target user u, the goal is to predict a preference order or a ranking (a special case of total orders of a set of alternatives) of these items according to u’s preferences.

Many approaches have been used the term *user preferences* for different purposes. In recommender systems, this term refers to user profiling, i.e., the way that users' tastes are *represented*, generally by means of a feature vector or a tensor. In general Artificial Intelligence (AI) research, this *user preferences* term refers to the preference *order* over objects or ranking inferred by a preference model. In our work we refer to *user preferences* as well as in AI research line: the *preference order* induced over objects.

Problem definition. Given a social network $\mathcal{N}$, a user u and a preference domain D, return a set P of order relations $\succ_u$ over D that describes u's preferences.

III. The Twitter Social Network

Through Twitter Streaming APIs[1], during the course of 95 days, we collected tweets related to Brazilian news. All tweets, retweets and quoted-status[2] containing some mention to the Brazilian newspaper Folha de São Paulo, whose Twitter user is *@folha*, were considered. In all, we collected 1,771,435 tweets and 292,310 distinct users in a time span of tweets posting times from Aug 7 2016 to Nov 9 2016.

Based on such tweets, we applied Latent Dirichlet Allocation (LDA) [3] to extract a set of topics to represent user's preferences. We got a total of 50 topics, that then were manually grouped in 7 more general topics. In the end, the preference domain is $D = \{politics, international, sports, entertainment, security, economy, others\}$. The same crawling and topic extraction strategies were used in [4], but for a short dataset (3 weeks time span).

IV. Mining User Preferences from Twitter

We compare two literature methods for preference mining [5], [6] based on implicit feedback and the baseline method given by explicit feedback. Despite proposed in literature, neither of them have been evaluated.

Favorites (FV). This is the method based on explicit feedback. Intuitively, we can build a preference ranking based on the number of favorites (likes) a user gives over a topic in preference domain. Though apparently straightforward, this method still face to the lack of ground truth problem as users not necessarily assign as favorite their preferred topics. Sometimes, favorite can be a strategy to store some important post which not necessarily is preferred.

Follower Network (FN). This method proposed in [5] is based on Twitter following relationship. The intuition is that if a user u_1 follows some personalities or celebrities, then u_1 prefers topics related to what those personalities represent. For instance, if u_1 follows some representative celebrity user u_2 from fashion world, and u_1 does not follow u_3 which is a religious leader, then u_1 prefers *fashion* topics over *religion* topics. Bringing to our context, we match each item in the preference domain D with specific Twitter users referring to publishing channels from Folha de S. Paulo. These users compose the set A. The match is defined by the function $f : A \rightarrow D$. For instance, topic $t = sports$ is assigned to user $u =$*@folhaesporte*, $t =$*politics* is $u =$*@folhapoder* and so on. Then, we built a follower/following network considering all users in our dataset and extract preferences of a given user u according to the following steps : (1) if u does not follow any $v \in A$ then $others \in D$ is preferred by u over all $o \in D - \{others\}$. Else, (2) for each $v \in A$ followed by u, add $t \succ o$ in P_{FN}, for $t, o \in D$ and $f(v) = t$. Remark that in this strategy we just have two levels of preferences: the most preferred objects and the others.

Topic Distribution (TD). This method was proposed in [6]. In this comparative study we slightly adapted it as our goal is not recommendation, just user profiling and preference extraction. The preference mining strategy is based on the number of tweets/retweets about some topic $t \in D$. The most tweeted/retweeted topic by u is the preferred one over the second most tweeted/retweeted which is preferred over the third one and so on. As example, if u posts three times about *politics* and two times about *sports*, then we can establish a preference order between *politics* and *sports* ($politics \succ sports$ and *sports* is preferred over the remaining topics $o \in D$). Here preference order levels can be deeper than in FN method.

In face of three different methods FV, FN and TD, and a common preference domain D, our proposal in this article lies at the combination of these methods in order to obtain a final set P_{GR} containing trusty preference relations for a given user, which can supply the lack of ground truth.

V. Preliminary Findings

Our goal in these preliminary experiments is to observe the preference set mined for each method for a given user. We seek to answer *how the methods FV, FN and TD overlap in their results?* We define the agreement score S_u for a given user u as

$$S_u = \frac{|P_{M_1} \cap P_{M_2} \cap ... \cap P_{M_n}|}{|P_{M_1} \cup P_{M_2} \cup ... \cup P_{M_n}|} \tag{1}$$

where P_{M_i} is the resultant preference set mined by method M_i and $n > 1$ is the total number of methods being combined. The higher S_u the higher the agreement among methods and thus, more reliable is the resultant preference ground truth set $P_{GT} = |P_{M_1} \cap P_{M_2} \cap ... \cap P_{M_n}|$.

Considering user u_1 (id=279635698), the mined preference sets are described below.

$P_{FV} = \{politics \succ_{u_1} international,\ politics \succ_{u_1} sports,\ politics \succ_{u_1} entertainment,\ politics \succ_{u_1} security,\ politics \succ_{u_1} economy,\ politics \succ_{u_1} others\}$

[1] https://dev.twitter.com/streaming/

[2] Quoted-status are retweets with comments

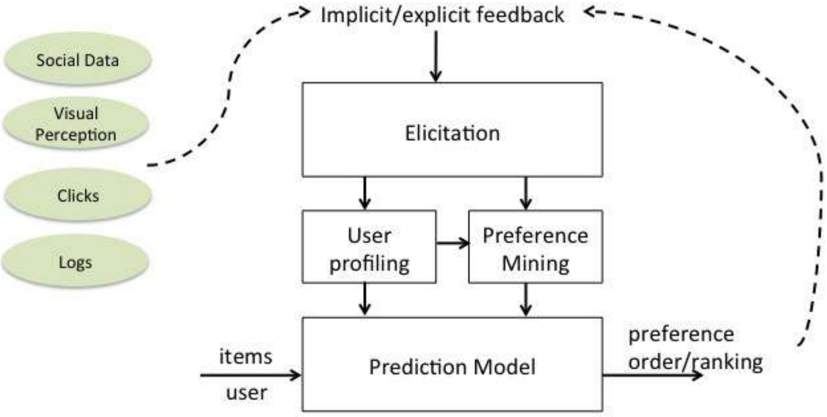


Fig. 1. Schema of a traditional preference prediction system.

$P_{FN} = \{politics \succ_{u_1} international, politics \succ_{u_1} entertainment, politics \succ_{u_1} security, politics \succ_{u_1} economy, politics \succ_{u_1} others, sports \succ_{u_1} international, sports \succ_{u_1} entertainment, sports \succ_{u_1} security, sports \succ_{u_1} economy, sports \succ_{u_1} others\}$

$P_{TD} = \{politics \succ_{u_1} international, politics \succ_{u_1} sports, politics \succ_{u_1} entertainment, politics \succ_{u_1} security, politics \succ_{u_1} economy, politics \succ_{u_1} others\}$

The corresponding agreement scores for u_1 are in Figure 2. Each $C_{M_1,\ldots,M_n}$ corresponds to the methods combination run.

Combination	S_{u_1}
$C_{FV,FN}$	0.454
$C_{FV,TD}$	1.0
$C_{FN,TD}$	0.454
$C_{FV,FN,TD}$	0.454

Fig. 2. Agreement scores among methods FV, FN and TD for user u_1

Combination	S_{avg}
$C_{FV,FN}$	0.061
$C_{FV,TD}$	0.429
$C_{FN,TD}$	0.066
$C_{FV,FN,TD}$	0.03

Fig. 3. Agreement scores among methods FV, FN and TD averaged for all users.

Figure 3 summarizes our results so far. The agreement score S_{avg} is the average of scores of all users in our dataset. Notice that the best score is for combination $C_{FV,TD}$. Also, FN is the worst performance, penalizing the full combination $C_{FV,FN,TD}$ score.

Discussions. There are other social network preference mining methods in literature not embraced in this study [7], [8]. The challenge is in applying methods proposed for very specific and diversified contexts in the same preference domain study. The technique from [7], for example, could not be applied in our news preference domain due to the lack of comparative sentences in tweets. In [8] the approach is ranking preference learning and the preferences are extracted from labels indicating fan page's political view in Facebook. A problem not tackled yet relies on consistency issues in resultant preference set P_{GT}. Given the transitive property of a preference relation $\succ$ over a domain D, P_{GT} is consistent if there is not any inferred preference $o \succ o \in P_{GT}$ for $o \in D$.

VI. Final Remarks

We have raised the discussion about evaluating without ground truth in social media research. In this context, we are studying the problem of preference mining in Twitter network. Three different existent methods have been implemented considering a common preference domain of news categories (sports, politics etc). In order to supply the lack of ground truth in our problem, we have proposed a combination strategy of the resultant set of preferences of each method to generate a final trustworthy set of user preferences. Our next steps will be study more methods [7], [8], [9] and social networks (Last.fm, Instagram, TikTok), and organize them according to their main features. Our final goal is to propose a framework able to extract preferences based on correlation and causality patterns, to eliminate the need of ground truth sets of preference relations. We expect that our method generalizes over other sources of data, for instance IoT domain [10] and web media [11].

References

[1] R. Zafarani and H. Liu, "Evaluation without ground truth in social media research," *Com. ACM*, vol. 58, no. 6, pp. 54–60, 2015.

[2] J. Furnkranz and E. Hullermeier, *Preference Learning*. Springer, New York, 2010.

[3] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, Mar. 2003.

[4] F. S. F. Pereira, S. de Amo, and J. Gama, "Detecting events in evolving social networks through node centrality analysis," *Large-scale Learning from Data Streams in Evolving Environments with ECML/PKDD*, 2016.

[5] ——, "On using temporal networks to analyze user preferences dynamics," in *Discovery Science: 19th International Conference, DS 2016, Bari, Italy, 2016.*, 2016.

[6] X. Liu, "Modeling users' dynamic preference for personalized recommendation," in *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI'15)*, 2015, pp. 1785–1791.

[7] F. S. F. Pereira and S. de Amo, "Mineracao de preferencias do usuario em textos de redes sociais usando sentencas comparativas," in *Symposium on Knowledge Discovery, Mining and Learning (KDMiLe)*, 2015, pp. 94–97.

[8] M. A. Abbasi, J. Tang, and H. Liu, "Scalable learning of users' preferences using networked data," in *Proceedings of the 25th ACM Conference on Hypertext and Social Media*, ser. HT '14. New York, NY, USA: ACM, 2014, pp. 4–12.

[9] H. Al-Jarrah, M. Al-Asa'd, S. A. Al-Zboon, S. K. Tawalbeh, M. M. Hammad, and M. AL-Smadi, "Resolving conflict of interests and recommending expert reviewers for academic publications using linked open data," in *2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS)*, 2019, pp. 91–98.

[10] T. Elsaleh, S. Enshaeifar, R. Rezvani, S. T. Acton, V. Janeiko, and M. Bermudez-Edo, "Iot-stream: A lightweight ontology for internet of things data streams and its use with data analytics and event detection services," *Sensors*, vol. 20, no. 4, 2020. [Online]. Available: https://www.mdpi.com/1424-8220/20/4/953

[11] T. R. Tangherlini, S. Shahsavari, B. Shahbazi, E. Ebrahimzadeh, and V. Roychowdhury, "An automated pipeline for the discovery of conspiracy and conspiracy theory narrative frameworks: Bridgegate, pizzagate and storytelling on the web," *PLOS ONE*, vol. 15, no. 6, pp. 1–39, 06 2020. [Online]. Available: https://doi.org/10.1371/journal.pone.0233879

6^{th} Workshop on Internet of Things—Enablers, Challenges and Applications

THE Internet of Things is a technology which is rapidly emerging the world. IoT applications include: smart city initiatives, wearable devices aimed to real-time health monitoring, smart homes and buildings, smart vehicles, environment monitoring, intelligent border protection, logistics support. The Internet of Things is a paradigm that assumes a pervasive presence in the environment of many smart things, including sensors, actuators, embedded systems and other similar devices. Widespread connectivity, getting cheaper smart devices and a great demand for data, testify to that the IoT will continue to grow by leaps and bounds. The business models of various industries are being redesigned on basis of the IoT paradigm. But the successful deployment of the IoT is conditioned by the progress in solving many problems. These issues are as the following:

- The integration of heterogeneous sensors and systems with different technologies taking account environmental constraints, and data confidentiality levels;
- Big challenges on information management for the applications of IoT in different fields (trustworthiness, provenance, privacy);
- Security challenges related to co-existence and interconnection of many IoT networks;
- Challenges related to reliability and dependability, especially when the IoT becomes the mission critical component;
- Zero-configuration or other convenient approaches to simplify the deployment and configuration of IoT and self-healing of IoT networks;
- Knowledge discovery, especially semantic and syntactical discovering of the information from data provided by IoT.

The IoT technical session is seeking original, high quality research papers related to such topics. The session will also solicit papers about current implementation efforts, research results, as well as position statements from industry and academia regarding applications of IoT. The focus areas will be, but not limited to, the challenges on networking and information management, security and ensuring privacy, logistics, situation awareness, and medical care.

Topics

The IoT session is seeking original, high quality research papers related to following topics:

- Future communication technologies (Future Internet; Wireless Sensor Networks; Web-services, 5G, 4G, LTE, LTE-Advanced; WLAN, WPAN; Small cell Networks…) for IoT,
- Intelligent Internet Communication,
- IoT Standards,
- Networking Technologies for IoT,
- Protocols and Algorithms for IoT,
- Self-Organization and Self-Healing of IoT Networks,
- Object Naming, Security and Privacy in the IoT Environment,
- Security Issues of IoT,
- Integration of Heterogeneous Networks, Sensors and Systems,
- Context Modeling, Reasoning and Context-aware Computing,
- Fault-Tolerant Networking for Content Dissemination,
- IoT Architecture Design, Interoperability and Technologies,
- Data or Power Management for IoT,
- Fog—Cloud Interactions and Enabling Protocols,
- Reliability and Dependability of mission critical IoT,
- Unmanned-Aerial-Vehicles (UAV) Platforms, Swarms and Networking,
- Data Analytics for IoT,
- Artificial Intelligence and IoT,
- Applications of IoT (Healthcare, Military, Logistics, Supply Chains, Agriculture, ...),
- E-commerce and IoT.

The session will also solicit papers about current implementation efforts, research results, as well as position statements from industry and academia regarding applications of IoT. Focus areas will be, but not limited to above mentioned topics.

Technical Session Chairs

- **Cao, Ning,** College of Information Engineering, Qingdao Binhai University
- **Chudzikiewicz, Jan,** Military University of Technology, Poland
- **Zieliński, Zbigniew,** Military University of Technology, Poland

Position Papers of the of the 17th Conference on Computer Science and Intelligence Systems pp. 71–78
DOI: 10.15439/2022F267
ISSN 2300-5963 ACSIS, Vol. 31

Distributed and Adaptive Edge-based AI Models for Sensor Networks (DAISeN)

Veselka Boeva, Emiliano Casalicchio
Shahrooz Abghari
Ahmed A. Al-Saedi, Vishnu Manasa Devagiri
Computer Science Department, Blekinge Institute of Technology
SE-371 79 Karlskrona, Sweden
Email: {vbx, emc}@bth.se

Andrej Petef
Peter Exner, Anders Isberg, Mirza Jasarevic
Sony Europe BV
R&D Center Europe
Lund, Sweden
Email: andrej.petef@sony.com

***Abstract*—This position paper describes the aims and preliminary results of the *Distributed and Adaptive Edge-based AI Models for Sensor Networks* (DAISeN)[1] project. The project ambition is to address today's edge AI challenges by developing advanced AI techniques that model knowledge from the sensor network and the environment to support the deployment of sustainable AI applications. We present one of the use cases being considered in DAISeN and review the state-of-the-art in three research domains related to the use case presented and directly falling into the project scope. We additionally outline the main challenges identified in each domain. The developed Global Navigation Satellite Systems (GNSS) activation model addressing the use case challenges is also briefly introduced. The future research studies planned for the remaining period of the project are finally outlined.**

I. Introduction

THE NUMBER of solutions that provide Artificial Intelligence (AI) and Machine Learning (ML) based systems has been growing recently. These solutions facilitate the creation of new smart products and services in many different fields. In addition, sensor networks are undergoing great expansion and development and the integration of AI and sensor networks benefits many areas such as Industry 4.0, healthcare, mobility, logistics, and many other Internet-of-Things (IoT) applications. However, this has also put new challenges in front of researchers and practitioners. New real-time AI and ML algorithms are needed along with different strategies to embed these algorithms in sensor boards and network nodes such as fog/edge nodes. For example, edge-based AI requires robust and adaptive models that take into account the temporal component of a data flow and allow for vertical and horizontal scaling of the decision-making process. These models must employ efficient learning algorithms that are capable of dealing with information varying over time and coping with large scale missing and inaccurate values. In addition, the decision-making models should be composable so that they can be distributed on the edge devices in order to ensure a trade-off between the decision accuracy, latency, and consumed energy per decision.

The IoT is an emerging key technology for future industries and the everyday lives of people, e.g., it has been playing an increasingly important role in healthcare, agriculture, home services, industrial processes, and transportation. Wireless Sensor Network (WSN) is an enabling technology for IoT [1], and, by definition, is the bridge between the physical world and the intelligence residing on the Internet. The integration of AI and sensor networks (by means of the IoT) are now realities that are changing our lives. Sensor networks are widely used to collect environmental parameters, e.g., in homes, buildings, and vehicles, where they are used as a source of information that aids the decision-making process and, in particular, it allows systems to learn and monitor activity. Although there are numerous advantages of sensor networks, it should be mentioned that they also consume energy and contribute to E-waste [2]. These place new stress on the environment and the smart world.

According to the World Economic Forum (WEF)[2], the IoT is undoubtedly one of the largest enablers for responsible digital transformation. WEF survey has outlined that more than 80% of IoT deployments are currently addressing, or have the potential to address the Sustainable Development Goals (SDGs)[3] defined by the United Nations, for example, industry, innovation, and infrastructure; smart cities and communities; affordable and clean energy (SDG #7); good health and well-being (SDG #3); clear water and sanitation (SDG #6); smart agriculture; responsible production and consumption (SDG #12).

Today's IoT solutions embed and leverage AI both in the end-user services and the network's management. To boost sustainability, IoT solutions need to be sustainable and usable. These goals are achievable only by means of advances in AI, decision making, and edge and fog computing. AI algorithms and decision-making models are at the core of state-of-the-art and future IoT applications and need to be distributed among IoT devices, such as WSN nodes, and edge and fog

This work is part of the Sony RAP 2020 Project "Distributed and Adaptive Edge-based AI Models for Sensor Networks"

[1] Daisen is a volcanic mountain located in Tottori Prefecture, Sanin Region of Japan.

[2] World Economic Forum, Internet-of-Things Guidelines for Sustainability (2018) http://www3.weforum.org/docs/IoTGuidelinesforSustainability.pdf

[3] United Nations' Sustainable Development Goals https://sdgs.un.org/goals

nodes, to scale vertically and horizontally, and to minimize energy consumption. At the same time, the management of complex IoT infrastructure (that includes fog WSN and edge/fog computing nodes) should be operated with energy-aware decision-making mechanisms that leverage distributed ML/AI techniques. The DAISeN project aims to address the challenges discussed above by developing:

- advanced AI algorithms for continual, shared, and evolving learning, that enable learning from multiple data sources by distributed training, and continual updating of the model;
- distributed decision-making models allowing vertical and horizontal scaling in order to guarantee high-quality decision-making concerning time consumption, energy consumption, and data communication.

This position paper is organized as in what follows. The research objectives of the DAISEeN project are described in Section II. In Section III-B we outline one of the use cases being considered in DAISeN and review the state-of-the-art ML and data mining techniques applicable at the edge. They are analyzed in relation to the use case to bring to evidence the challenges and gaps the project aims to fill. Preliminary results are reported in Section IV. Finally, the paper is concluded by our outlook (Section V).

II. Research Objectives

The main focus of the proposed research lies in the usability of AI on edge devices and fog nodes to improve the performance and sustainability of sensor networks and the training process.

The DAISEeN project will investigate methods for transferring and adapting AI algorithms to edge devices with limited computational performance. The ambition is the development of resource-aware AI algorithms that can be run on edge nodes. These algorithms take into account the hardware and software capabilities of the edge nodes and the capabilities of the communication links between these nodes, always keeping in mind that a balance is found between the limited energy resources of edge devices and the complexity of the AI algorithms. Therefore, the network traffic between the edge devices has to be kept low. That can be achieved by several methods such as context-aware techniques [3], dynamic device clustering, role assignments, or intelligent sensor fusion, and data reasoning techniques that can account for dynamically changing surrounding environment, including context prediction.

Another challenge addressed by DIASeN is the development of advanced AI algorithms for continual, shared, and evolving learning that enable learning from multiple data sources by distributed training and continual updating of the model. This can be achieved by developing unsupervised and semi-supervised methods to automate knowledge extraction and learning in data stream scenarios [4], [5]. The main problem investigated is how the newly arrived information can be taken into account in the learning phase and can be used for continuous adaptation of the learned model [6]. In addition, it will be studied how to develop, train, and evaluate a model with no direct access to labeled data. Candidate approaches to address those challenges are:

1) dynamic unsupervised and semi-supervised learning models that are robust to the appearance of drifting context and additionally enable to learn from multiple data sources by distributed training, and continual updating and evolving of the model [4], [6], [7], [8];
2) development of dynamic techniques for automatic annotation (labeling) of the data;
3) usage of transfer learning techniques enabling reuse of knowledge from training in earlier tasks to subsequent tasks.

The other research ambition of DAISeN is the design of distributed/composable data mining models. These models will allow to vertically/horizontally scale the decision making in order to guarantee high quality decisions in edge computing environments. This can be achieved by adopting ML models that can predict the computational level (i.e., cloud vs. edge) with respect to the network operational context (e.g., latency vs. accuracy). Such models will allow one to run the lighter but less accurate models at the edge for the sake of latency, and the computation-intensive but higher-accuracy models in the cloud. The focus is on developing data mining and ML techniques to maintain local models embedded in the edge devices and further integrate low-level edge devices' observations into a global model [4], [5], [9]. Computed at a higher level, the latest can produce reliable decisions based on the available input data. Those techniques will produce a global model even when data from some devices are missing due to network changes or degradation.

III. DAISeN Use Case and State of the Art

A. A context-aware GNSS activation use case

Sony provides software solutions in smart logistics for monitoring and tracking goods. GNSS is the used positioning technique for detecting the tracker's current position. GNSS is known to perform well in open sky environments. However, the trackers may be in any place, such as open outdoors, crowded city areas, indoors, and so on. Sony would like to perform context-aware control of GNSS activation by automatically and accurately detecting indoor/outdoor localization of the trackers by consuming the least energy. The Sony requirement is to use radio signals received from Long-Term Evolution (LTE) base stations to detect the environment (indoor/outdoor). The main idea is that the propagation of radio signals is affected by the environment. Different environmental scenarios have different signal strength characteristics. By learning different signal strength characteristics, it would be possible to determine the tracker's environment.

The setup explained above requires advanced AI solutions that can detect indoor/outdoor localization (environmental context) of the tracker for controlling GNSS activation in order to save power. In addition, these models are expected to be able to continuously adapt to new scenarios and environments, as well as to learn in the distributed framework from many trackers to improve environmental context detection.

B. *Current state of the art and challenges*

In this section, we review ML and data mining techniques related to the use case described in Section III-A. The discussed techniques fall into three main research domains: context-awareness at the edge, continual and evolving learning, and federated learning, see Table I.

1) Context awareness at the edge: Our considerations are limited to indoor/outdoor context-awareness methods at the edge due to the fact that it is related to the use case discussed above (see Section III-A).

Efficient knowledge discovery is critical for the optimized operation and management of IoT/sensor networks. Context may affect the complex systems' operation and management procedures at various levels, from the physical to the communication, up to the application level [10]. For example, as discussed above, positioning edge devices, such as smartphones and trackers, in outdoor areas typically rely on GNSS such as the Global Positioning System (GPS), which performs well in open sky environments. However, these devices may be in any place, such as deep indoors, metal containers, crowded urban areas, etc. In addition, the GPS consumes too much energy to be useful for many applications. Therefore, detecting indoor/outdoor and providing this context-aware information in various environments may be helpful and lead to battery-saving solutions. Many indoor/outdoor detection methods have already been proposed.

In [11], these methods are classified into two main groups: threshold-based techniques and ML-based techniques. Approaches in the first group use fixed detection rules and thresholds, such as a sensor reading above a certain value, to classify an edge node state (e.g., indoor/outdoor). The second group of solutions uses ML algorithms to detect indoor/outdoor status based on features extracted from smartphones, edge nodes, and in general, embedded sensors. ML-based indoor/outdoor environment detection techniques are in the focus of our interest and are reviewed further in this section.

In [12], an approach, entitled SenseMe, uses the C4.5 algorithm on data generated from GPS, gyroscope, accelerometer, and the Bluetooth module to sense environmental context and the context-aware location. In [13], the authors propose a sound-based indoor/outdoor detection method that utilizes binary classification of the environment's acoustic reverberation features. Canovas et al. [14] employ a binary classification technique on the Received Signal Strength Indicator (RSSI) from 802.11 access points to identify a pedestrian's indoor or outdoor status. Ashraf et al. [15] propose MagIO, a solution that utilizes magnetic field signals sensed by smartphones for detecting indoor/outdoor states. Magnetic field features are classified with different ML algorithms, including Support Vector Machines (SVM), Gradient Boosting Machines (GBM), Random Forest (RF), k-Nearest Neighbor (kNN), and Decision Trees (DT). In [16], the authors apply an ML algorithm to classify the neighboring GSM station's signal in different environments and identify the users' current context by signal recognition. Radu et al. [17] propose to detect indoor/outdoor context by employing co-training according to the feature of light, magnetic, and cell sensors. The proposed solution can automatically learn characteristics of new environments and devices, thereby providing a detection accuracy exceeding 90% even in unfamiliar circumstances. Multiple contextual features are also used in [18], which leveraged J48 and other ML algorithms to detect the indoor/outdoor state with high accuracy. An interesting hybrid solution that integrates unsupervised and supervised algorithms relying on the location accuracy and signal strength is introduced in [19].

Challenges: Most of the reviewed approaches rely on the presence of a large amount of labeled data and report higher performance on datasets coming from the same locations/devices as those used to build the model than on new environments. However, when detecting environmental (indoor/outdoor) context at the edge level in real-world scenarios usually labeled data is scarce or entirely missing. Furthermore, the environment dynamic and the context complexity should be taken into account, but at the same time, keeping in mind that the detection models should be light in order to be able to run on the device [3]. Evidently, novel context-aware data mining and learning techniques are needed. These must be resource-efficient, but also should be able to support continual learning from multiple sources and robust model adaptation to new environments.

2) Continual and evolving learning: The main ideas depicted in the continual learning paradigm are knowledge sharing, adaptation, and transfer [20]. Continual learning algorithms may have to deal with catastrophic forgetting [21], data distribution shifts [22], or imbalanced or scarce data problems [23]. Catastrophic forgetting [21] refers to a model experiencing performance degradation at previously learned concepts when trained sequentially in learning new concepts. The catastrophic forgetting is a significant challenge to tackle in the continual learning context since, by definition, the continual learning setting deals with sequences of classes or tasks. Other challenges that should be considered are data distribution shifts and the emergence of new classes. Changes in the data distribution over time are commonly referred to as concept drift. Gepperth and Hammer [22] define two kinds of concept drift: virtual and real. Virtual concept drift concerns the input distribution and may due to imbalanced classes over time. On the contrary, real concept drift is caused by novelty on data or new classes and can be detected by its effect on, e.g., classification accuracy. The continual learning model has to detect the change and automatically fix it. An undetected shift in the data distribution will lead to forgetting. Online change detection algorithms deal with this challenge as it is shown in [24], [25].

The study [20] surveys different supervised continual learning approaches and classifies them into three main categories: replay, regularization-based, and parameter isolation methods. This classification is based on how task-specific information is stored and used throughout the sequential learning process.

TABLE I
2.2. DISTRIBUTED AI REQUIREMENTS IN RELEVANT STATE-OF-THE-ART AREAS

Domain	**State-of-the-art area**	**Relevant use case requirements**
Computations at the edge	Context awarness at the edge	Detect indoor/outdoor localization of the edge node (tracker)
Learning from streaming data	Continual and evolving learning	Continuously update the model when new data are available
Distributed AI	Federated learning	Distributed learning from many edge nodes (trackers) for improving the context detection

The most important studies published in these three categories are summarized in Table II. Replay methods replay previous task samples while learning a new task to alleviate forgetting. The replayed samples are either reused as inputs for rehearsal or for constrained optimization of the new task loss to prevent previous task interference. Rehearsal methods explicitly retrain a limited subset of stored samples while training on new tasks [26], [27], [28], [29]. In the absence of previous samples, pseudo rehearsal is an alternative strategy used in early works with shallow neural networks [30], [31], [32], [33]. Constrained optimization is considered an alternative solution to rehearsal by leaving more freedom for backward/-forward transfer. Rehearsal might be prone to overfitting the subset of stored samples and appears to be bounded by joint training [34], [35]. Regularization-based methods introduce an additional term in the loss function, consolidating previous knowledge when learning new data. These methods can further be divided into data-focused and prior-focused methods. Knowledge distillation from a previous model to the trained model on the new data is the primary building block in data-focused methods [36], [37], [38], [39]. Prior-focused methods mitigate forgetting by estimating a distribution over the model parameters used prior when learning from new data [40], [41], [42], [43], [44], [45]. Parameter isolation methods isolate parameters for specific tasks and can guarantee maximal stability by fixing the parameter subsets of previous tasks. For example, Mallya and Lazebnik [46] have proposed an approach that uses weight-based pruning techniques to free up redundant parameters across all layers of a deep network after it has been trained for a task. Another approach built upon ideas from fixed network quantization and pruning is introduced in [47]. A different approach for continual learning is proposed in [48], namely, it searches for the best neural architecture for each coming task via sophisticated reinforcement learning strategies. The studies in [49], [50] also fall into the category of dynamic architecture-based continual learning solutions.

Most existing continual learning approaches are designed in a supervised fashion assuming all data from new tasks have been manually annotated. However, in many real-world applications of continual learning, e.g., learning from sensor data streams to make real-time classification, the availability of relevant labeled data is often low or even non-existing [51], [52], [53]. Most real-world data is usually not consistently labeled, i.e., there is no explicit indication of the exact periods of relevant events and occurrences of interesting trends, which breaks down the traditional supervised learning paradigm. Data labeling is mostly done manually by human experts. This process is, however, labor-intensive, time-consuming, and very expensive. Unsupervised continual learning, which is expected to tackle the aforementioned issues, has not been well studied [54]. Caron et al. [55] have proposed to iteratively cluster features and update the model with subsequently assigned pseudo labels obtained by applying a standard clustering algorithm. Another recent work proposes to perform clustering and model updates simultaneously to address the model's instability during the training phase [56]. However, these methods only work on static datasets and cannot learn new knowledge incrementally. In [57], the authors introduced a simple and effective method that, in an unsupervised setting, can be adapted to existing supervised continual learning approaches. The authors propose to use a pseudo label instead of the ground truth to make continual learning feasible in unsupervised mode. The pseudo labels of new data are obtained by applying a global clustering algorithm.

Evolving clustering models are good candidates to tackle concept drift scenarios. They have been designed to mine massive datasets or online continuous data streams in an unsupervised learning context by grouping and by summarizing data in a fast-incremental manner. Evolving clustering methods can process data stepwise and update and evolve cluster partitions in incremental learning steps [58], [59]. According to [58], different phases of an evolutionary clustering algorithm can be categorized into matching, accommodating new data, and model refinement. Dynamic clustering is also a form of online/incremental unsupervised learning [4], [6], [7], [9], [60]. However, it considers the incremental fashion of building the clustering model and self-adaptation of the built model. Dynamic clustering algorithms can split or merge the clusters based on the need.

Challenges: Data is often collected from unreliable sources, possibly having missing values and inaccurate labels. Hence, there is a need for a conceptually new learning framework to support continual and evolving learning under uncertainty and noise [51], [53]. In general, to take advantage of new developments in AI research, such as shared and continual learning [61], we need novel data mining and learning models. Those models should be capable of dealing with unlabeled data having large-scale missing and inaccurate labels, enabling learning from multiple data sources via distributed training and continual evolution of the model [62], [63] while efficiently

TABLE II
MAIN CATEGORIES OF SUPERVISED CONTINUAL LEARNING METHODS ACCORDING TO THE SURVEY PUBLISHED IN [20].

Method's category	Sub-categories	Studies	Pros & Cons
Replay	Rehearsal	[26], [27], [28], [29]	Limited scalability, Privacy issues, No clear policy for unbalanced tasks, Task-agnostic
	Pseudo-rehearsal	[30], [31], [32], [33]	
	Constrained	[34], [35]	
Regularization-based	Prior-focused	[40], [41], [42], [43], [44], [45]	Prioritizing privacy, Alleviated memory requirements, Task-agnostic
	Data-focused	[36], [37], [38], [39]	
Parameter isolation	Fixed Network	[47], [46]	Efficient memory, Prevents scalable class incremental setup
	Dynamic Architectures	[48], [49], [50]	

dealing with catastrophic forgetting and automatically adapting to real concept drift.

3) Federated learning: Federated learning (FL) has been introduced as promising collaborative learning, where edge devices such as smartphones, tablets, sensors, etc. keep their local data in their premises and exchange model parameters with a central server for global model aggregation [64], [65]. The global model is updated by averaging the local model parameters received by all the edge devices and is shared with them again. These operations are repeated at each iteration round. This setup has many advantages but also challenges such as expensive communication, systems heterogeneity due to the verity of devices in federated networks, and privacy concerns [66]. The iterative nature of FL requires massive communication between the central server and edge devices to train a global model [65]. The communication overhead at each iteration is not negligible, especially for complex models, large-scale applications, and high-frequency updates, and it becomes a challenge to be addressed [64], [65], [67]. Recently, many studies aiming to reduce communication costs have been proposed. For example, [68] use models of different sizes to address heterogeneous clients equipped with different computation and communication capabilities, while the work in [69] uses decentralized collaborative learning in combination with the master-slave model. The majority of solutions that address the problem of reducing network overhead in FL could be classified into two main categories. The first category incorporates works that reduce the total number of bits transferred for each local update through data compression. The second category includes studies that aim at reducing the number of local updates during the training process.

The authors of [70] propose an enhanced FL technique by introducing an asynchronous learning strategy on the clients and a temporally weighted aggregation of the local models on the server. The layers of the deep neural networks are categorized into shallow and deep layers. The parameters of the deep layers are updated less frequently than those of the shallow layers. In addition, a temporally weighted aggregation strategy is applied on the server to make use of the previously trained local models, thereby enhancing the accuracy and convergence of the central model. The paper [71] designs two novel strategies to reduce communication costs. The first relies on lossy compression on the global model sent server-to-client. The second strategy uses Federated Dropout (FD), which allows users to efficiently train locally on smaller subsets of the global model and reduces the client-to-server communication and the local computation. Deep Gradient Compression (DGC) is proposed to significantly reduce communication bandwidth [72]. The authors of [73] introduce a new compression framework, entitled Sparse Ternary Compression, that is specifically designed to meet the requirements of the FL environment. The authors of [74] implement a Federated Optimisation (FedOpt) approach by designing a novel compression algorithm for efficient communication. Then, they integrate additively homomorphic encryption with differential privacy to prevent data from being leaked. Malekijoo et al. [75] develop a novel framework that significantly decreases the size of updates while transferring weights from the deep learning model between clients and their servers. A novel algorithm, namely FetchSGD, that compresses model updates using a Count Sketch, and takes advantage of the mergeability of sketches to combine model updates from many workers, is proposed by [76]. Xu et al. [77] present a Federated Trained Ternary Quantization (FTTQ) algorithm, which optimizes the quantized networks on the clients through a self-learning quantization factor.

A novel FedMed method with adaptive aggregation is proposed using the topK strategy to select the top workers with the lowest losses to update the model parameters in each round in [74]. Asad et al. [78] have provided a novel filtering procedure on each local update that only transfers significant gradients to the server. The study proposed by [79] identifies the relevant updates of participants and uploads them only to the server. Specifically, at each round, the participants receive the global tendency and check the relevancy of their local updates with the global model. If they align, the updates are uploaded. An FL two-step client selection protocol based on resource constraints instead of the random client selection is proposed by [80]. FedPSO, a global model update algorithm, transmits the model weights only to the client that has provided the best score (such as accuracy or loss) to the cloud server [81].

Challenges: So far there is no evidence of how FL approaches are reducing the number of bits transferred compared to FL approaches that reduce the number of local updates. However, concerning the latter category of approaches, it is vital to find out more efficient FL schemes other than FedAvg, which converge with the same speed as FedAvg and apply to any FL applications [82]. For example, the studies in [83] and [84] have explored an approach that applies clustering optimization to bring efficiency and robustness in FL's communication: only the most representative updates are uploaded to the central server for reducing network communication costs.

IV. Preliminary Results

A. An inductive system monitoring approach for GNSS activation

In order to address the above challenges, we have designed a GNSS component activation model for mobile tracking devices which automatically detects indoor/outdoor environments using the radio signals received from LTE base stations [3]. We use an Inductive System Monitoring (ISM) technique [85] to model environmental scenarios captured by each tracker via extracting clusters of corresponding value ranges from base stations' signal strength. The ISM-based model is built by using the tracker's historical data labeled with GPS coordinates. The built model is further refined by applying it to the data without GPS location collected by the same device. This procedure allows us to identify the clusters that describe semi-outdoor scenarios. Thus, the model enables to discriminate between two outdoor environmental categories: open outdoor and semi-outdoor. Each cluster models an open outdoor or a semi-outdoor scenario by defining a range of allowable values for each base station in a given input vector. The vector of high values and the vector of low values in a cluster are considered as the cluster's representatives describing a specific environmental scenario. Evidently, the proposed model supplies the user with easily interpretable representations of the device's outdoor environmental scenarios. Note that the built ISM-based model does not contain the description of the indoor environmental scenarios, i.e., during the monitoring phase, data samples that do not fit any of the clusters are interpreted as belonging to the indoor environment. As a result, the built model is small and has modest requirements with respect to storage and computations.

B. Evaluation results

The proposed ISM-based GNSS activation approach is studied and evaluated on real-world data provided by Sony [3]. The used dataset contains radio signal measurements collected by five trackers and their geographical location in various environmental scenarios. We have explored the performance of the built ISM-based GNSS component activation model on this dataset in three different experiments. The obtained results have been analyzed and interesting patterns about the GNSS activation problem have been extracted. For example, we have conducted an experiment in which we use data with GPS coverage collected by each tracker to build a model representing the tracker's behavior. In addition, either the collected signals strengths by the trackers have been sorted in descending order, or they have been left as initially received. For testing, data without GPS coverage have been used. Table III lists the accuracy of the models for each tracker device with and without sorting the signal strengths. As one can observe, most models (M-di, $i = 1, \ldots, 5$) exhibit, except d5, higher accuracy in the case of sorted signal strengths. In addition, in another experiment, we have discovered that the models built on unshuffled data have shown higher performance. Furthermore, we have compared the performance of the model built on the data collected from all five devices with that of the individual trackers' models. The latter have demonstrated higher performance than the overall model. Evidently, the use of models with sorted and unshuffled signals is recommended. In addition, the customization of each tracker's model to the device specific environmental scenarios is preferred, since it ensures higher performance.

TABLE III
Model's accuracy (%) on data without GPS coverage for sorted and unsorted trackers' signal strengths

Model	Sorted signals	Unsorted signals
M-d1	99.89	64.44
M-d2	57.15	49.82
M-d3	61.49	60.40
M-d4	68.81	56.04
M-d5	71.94	72.83

The obtained evaluation results will be used for further improvement and optimization of the developed model (see our future plans in Section V). The company is currently evaluating and testing the model in the field.

V. Outlook

This paper describes the main objectives, identified challenges, and preliminary results of the DAISeN project. The main findings, valid for the reviewed research domains falling into the scope of DAISeN, reveal that in order to address the current challenges at the edge, we need novel resource and energy-efficient data mining algorithms and ML models robust to noisy, unlabeled, and missing data. Additionally, algorithms that enable learning from multiple data sources by distributed training and continual model adaptation are required.

In order to address the identified challenges, in the first half of the project, we have developed a novel GNSS component activation model for mobile tracking devices which is able to automatically detect indoor/outdoor environments based on the radio signals received from LTE base stations. The future research studies planned for the remaining period of the project involve the development of a domain integration GNSS activation technique that enables the integration of GNSS activation models built on different domains (devices/locations) into an overall model. In addition, we have the ambition to design a distributed GNSS activation framework that is enabled to create a shared model with the help of a large number of edge devices.

References

[1] J. A. Manrique *et al.*, "Contrasting internet of things and wireless sensor network from a conceptual overview," in *2016 IEEE Int. Conference on Internet of Things and IEEE Green Computing and Communications and IEEE Cyber, Physical and Social Computing and IEEE Smart Data*, pp. 252–257. [Online]. Available: https://doi.org/10.1109/iThings-GreenCom-CPSCom-SmartData.2016.66

[2] R. Arshad *et al.*, "Green iot: An investigation on energy saving practices for 2020 and beyond," *IEEE Access*, vol. 5, pp. 15 667–15 681, 2017. [Online]. Available: https://doi.org/10.1109/ACCESS.2017.2686092

[3] S. Abghari, V. Boeva, E. Casalicchio, and P. Exner, "An inductive system health monitoring approach for gnss activation," in *Artificial Intelligence Applications and Innovations*. Springer Nature Switzerland, 2022. [Online]. Available: https://doi.org/10.1007/978-3-031-08337-2_36

[4] V. M. Devagiri, V. Boeva, and E. Tsiporkova, "Split-merge evolutionary clustering for multi-view streaming data," *Procedia Computer Science*, vol. 176, pp. 460–469, 2020. [Online]. Available: https://doi.org/10.1016/j.procs.2020.08.048

[5] C. Åleskog, V. M. Devagiri, and V. Boeva, *A Graph-Based Multi-view Clustering Approach for Continuous Pattern Mining*. Cham: Springer International Publishing, 2022, pp. 201–237. [Online]. Available: https://doi.org/10.1007/978-3-030-95239-6_8

[6] V. Boeva *et al.*, "Bipartite split-merge evolutionary clustering," in *Int. conference on agents and AI*. Springer, 2019, pp. 204–223. [Online]. Available: https://doi.org/10.1007/978-3-030-37494-5_11

[7] E. Lughofer, "A dynamic split-and-merge approach for evolving cluster models," *Evolving systems*, vol. 3, no. 3, pp. 135–151, 2012. [Online]. Available: https://doi.org/10.1007/s12530-012-9046-5

[8] C. Nordahl, V. Boeva, G. Håkan, and M. P. Netz, "Evolvecluster: an evolutionary clustering algorithm for streaming data," *Evolving Systems*. [Online]. Available: https://doi.org/10.1007/s12530-021-09408-y

[9] V. M. Devagiri, V. Boeva, and S. Abghari, "A multi-view clustering approach for analysis of streaming data," in *AI Applications and Innovations*, I. Maglogiannis, J. Macintyre, and L. Iliadis, Eds. Springer International Publishing, 2021, pp. 169–183. [Online]. Available: https://doi.org/10.1007/978-3-030-79150-6_14

[10] T. E. Bogale *et al.*, "Machine intelligence techniques for next-generation context-aware wireless networks," *Int. Telecommunication Union Journal*, 2018.

[11] Y. Zhu *et al.*, "A fast indoor/outdoor transition detection algorithm based on machine learning," *Sensors*, vol. 19, no. 4, p. 786, 2019. [Online]. Available: https://doi.org/10.3390/s19040786

[12] P. Bhargava *et al.*, "Senseme: a system for continuous, on-device, and multi-dimensional context and activity recognition," in *Proceedings of the 11th Int. Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*, 2014, pp. 40–49. [Online]. Available: http://dx.doi.org/10.4108/icst.mobiquitous.2014.257654

[13] R. Sung *et al.*, "Sound based indoor and outdoor environment detection for seamless positioning handover," *ICT Express*, vol. 1, no. 3, pp. 106–109, 2015. [Online]. Available: https://doi.org/10.1016/j.icte.2016.02.001

[14] O. Canovas *et al.*, "Wifiboost: A terminal-based method for detection of indoor/outdoor places," in *Proceedings of the 11th Int. Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*, 2014, pp. 352–353. [Online]. Available: https://doi.org/10.4108/icst.mobiquitous.2014.258063

[15] I. Ashraf *et al.*, "Magio: Magnetic field strength based indoor- outdoor detection with a commercial smartphone," *Micromachines*, vol. 9, no. 10, 2018. [Online]. Available: https://doi.org/10.3390/mi9100534

[16] W. Wang *et al.*, "Indoor-outdoor detection using a smart phone sensor," *Sensors*, vol. 16, no. 10, p. 1563, 2016. [Online]. Available: https://doi.org/10.3390/s16101563

[17] V. Radu *et al.*, "A semi-supervised learning approach for robust indoor-outdoor detection with smartphones," in *Proceedings of the 12th ACM Conf. on Embedded Network Sensor Systems*, 2014, p. 280–294. [Online]. Available: https://doi.org/10.1145/2668332.2668347

[18] T. Anagnostopoulos *et al.*, "Environmental exposure assessment using indoor/outdoor detection on smartphones," *Personal and Ubiquitous Computing*, vol. 21, no. 4, pp. 761–773, 2017. [Online]. Available: https://doi.org/10.1007/s00779-017-1028-y

[19] R. P. Souza *et al.*, "A big data-driven hybrid solution to the indoor-outdoor detection problem," *Big Data Research*, vol. 24, p. 100194, 2021. [Online]. Available: https://doi.org/10.1016/j.bdr.2021.100194

[20] M. D. Lange *et al.*, "A continual learning survey: Defying forgetting in classification tasks," *IEEE transactions on pattern analysis and machine intelligence*, vol. PP, 2021. [Online]. Available: https://doi.org/10.1109/TPAMI.2021.3057446

[21] R. M. French, "Catastrophic forgetting in connectionist networks," *Trends in cognitive sciences*, vol. 3, no. 4, pp. 128–135, 1999. [Online]. Available: https://doi.org/10.1016/S1364-6613(99)01294-2

[22] A. Gepperth and B. Hammer, "Incremental learning algorithms and applications," in *European symposium on artificial neural networks (ESANN)*, 2016. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01418129

[23] P. Sprechmann *et al.*, "Memory-based parameter adaptation," in *Int. Conference on Learning Representations*, 2018. [Online]. Available: https://openreview.net/forum?id=rkfOvGbCW

[24] V. Moens and A. Zénon, "Learning and forgetting using reinforced bayesian change detection," *PLoS computational biology*, vol. 15, no. 4, p. e1006713, 2019. [Online]. Available: https://doi.org/10.1371/journal.pcbi.1006713

[25] Y. Sun *et al.*, "Planning to be surprised: Optimal bayesian exploration in dynamic environments," in *Int. conf. on AGI*. Springer, 2011, pp. 41–51. [Online]. Available: https://doi.org/10.1007/978-3-642-22887-2_5

[26] M. De Lange and T. Tuytelaars, "Continual prototype evolution: Learning online from non-stationary data streams," in *Proc. of the IEEE/CVF Int. Conf. on Comp. Vision*, 2021, pp. 8250–8259. [Online]. Available: https://doi.org/10.1109/iccv48922.2021.00814

[27] S.-A. Rebuffi *et al.*, "icarl: Incremental classifier and representation learning," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 2001–2010. [Online]. Available: https://doi.org/10.1109/CVPR.2017.587

[28] D. Isele and A. Cosgun, "Selective experience replay for lifelong learning," in *Proc. of the AAAI Conference on AI*, vol. 32, no. 1, 2018.

[29] D. Rolnick *et al.*, "Experience replay for continual learning," *Advances in Neural Information Processing Systems*, vol. 32, 2019.

[30] H. Shin *et al.*, "Continual learning with deep generative replay," *Advances in neural information processing systems*, vol. 30, 2017. [Online]. Available: https://dl.acm.org/doi/10.5555/3294996.3295059

[31] F. Lavda *et al.*, "Continual classification learning using generative models," *Proceedings of the 32nd Conference on Neural Information Processing Systems (NeurIPS) 2018*, 2018.

[32] J. Ramapuram *et al.*, "Lifelong generative modeling," *Neurocomputing*, vol. 404, pp. 381–400, 2020. [Online]. Available: https://doi.org/10.1016/j.neucom.2020.02.115

[33] C. Atkinson *et al.*, "Pseudo-rehearsal: Achieving deep reinforcement learning without catastrophic forgetting," *Neurocomputing*, vol. 428, pp. 291–307, 2021. [Online]. Available: https://doi.org/10.1016/j.neucom.2020.11.050

[34] D. Lopez-Paz and M. Ranzato, "Gradient episodic memory for continual learning," *Advances in neural information processing systems*, vol. 30, 2017. [Online]. Available: https://dl.acm.org/doi/10.5555/3295222.3295393

[35] A. Chaudhry *et al.*, "Riemannian walk for incremental learning: Understanding forgetting and intransigence," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 532–547. [Online]. Available: https://doi.org/10.1007/978-3-030-01252-6_33

[36] Z. Li and D. Hoiem, "Learning without forgetting," *IEEE transactions on pattern analysis and machine intelligence*, vol. 40, no. 12, pp. 2935–2947, 2017. [Online]. Available: https://doi.org/10.1109/TPAMI.2017.2773081

[37] H. Jung *et al.*, "Less-forgetting learning in deep neural networks," *Proceedings of the AAAI Conference on Artificial Intelligence*, no. 1, 2018. [Online]. Available: https://doi.org/10.1609/aaai.v32i1.11769

[38] A. Rannen *et al.*, "Encoder based lifelong learning," in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 1320–1328. [Online]. Available: https://doi.org/10.1109/ICCV.2017.148

[39] J. Zhang *et al.*, "Class-incremental learning via deep model consolidation," in *Proc. of the IEEE/CVF WACV*, 2020, pp. 1131–1140. [Online]. Available: https://doi.org/10.1109/WACV45572.2020.9093365

[40] J. Kirkpatrick *et al.*, "Overcoming catastrophic forgetting in neural networks," *Proceedings of the national academy of sciences*, vol. 114, no. 13, pp. 3521–3526, 2017. [Online]. Available: https://doi.org/10.1073/pnas.1611835114

[41] S.-W. Lee *et al.*, "Overcoming catastrophic forgetting by incremental moment matching," *Advances in neural information processing systems*, vol. 30, 2017. [Online]. Available: https://dl.acm.org/doi/10.5555/3294996.3295218

[42] F. Zenke, B. Poole, and S. Ganguli, "Continual learning through synaptic intelligence," in *International Conference on Machine Learning*. PMLR, 2017, pp. 3987–3995.

[43] X. Liu *et al.*, "Rotate your networks: Better weight consolidation and less catastrophic forgetting," in *2018 24th ICPR*. IEEE, 2018, pp. 2262–2268. [Online]. Available: https://doi.org/10.1109/ICPR.2018.8545895

[44] R. Aljundi *et al.*, "Memory aware synapses: Learning what (not) to forget," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 139–154.

[45] A. Chaudhry *et al.*, "Efficient lifelong learning with a-GEM," in *Int. Conf. on Learning Representations*, 2019.

[46] A. Mallya and S. Lazebnik, "Packnet: Adding multiple tasks to a single network by iterative pruning," in *Proc. of IEEE CVPR*, 2018, pp. 7765–7773. [Online]. Available: https://doi.org/10.1109/CVPR.2018.00810

[47] A. Mallya *et al.*, "Piggyback: Adapting a single network to multiple tasks by learning to mask weights," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 67–82.

[48] J. Xu and Z. Zhu, "Reinforced continual learning," *Advances in Neural Information Processing Systems*, vol. 31, 2018. [Online]. Available: https://dl.acm.org/doi/10.5555/3326943.3327027

[49] R. Aljundi *et al.*, "Expert gate: Lifelong learning with a network of experts," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 3366–3375. [Online]. Available: https://doi.org/10.1109/CVPR.2017.753

[50] A. Rosenfeld and J. K. Tsotsos, "Incremental learning through deep adaptation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 42, no. 3, pp. 651–663, 2018. [Online]. Available: https://doi.org/10.1109/TPAMI.2018.2884462

[51] Z. Chen and B. Liu, "Lifelong machine learning," *Synthesis Lectures on AI and ML*, vol. 12, no. 3, pp. 1–207, 2018.

[52] G. De Francisci Morales *et al.*, "Iot big data stream mining," in *Proceedings of the 22nd ACM SIGKDD int. conference on knowledge discovery and data mining*, 2016, pp. 2119–2120. [Online]. Available: https://doi.org/10.1145/2939672.2945385

[53] H. M. Gomes *et al.*, "Machine learning for streaming data: state of the art, challenges, and opportunities," *ACM SIGKDD Explorations Newsletter*, vol. 21, no. 2, pp. 6–22, 2019. [Online]. Available: https://doi.org/10.1145/3373464.3373470

[54] M. Masana *et al.*, "Class-incremental learning: survey and performance evaluation on image classification," *arXiv preprint arXiv:2010.15277*, 2020.

[55] M. Caron, Bojanowski *et al.*, "Deep clustering for unsupervised learning of visual features," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 132–149. [Online]. Available: https://doi.org/10.1007/978-3-030-01264-9_9

[56] X. Zhan *et al.*, "Online deep clustering for unsupervised representation learning," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 6688–6697. [Online]. Available: https://doi.org/10.1109/cvpr42600.2020.00672

[57] J. He and F. Zhu, "Unsupervised continual learning via pseudo labels," *arXiv preprint arXiv:2104.07164*, 2021.

[58] A. Bouchachia, "Evolving clustering: An asset for evolving systems," *IEEE SMC Newsletter*, vol. 36, pp. 1–6, 2011.

[59] M. Zopf *et al.*, "Sequential clustering and contextual importance measures for incremental update summarization," in *Proceedings of COLING 2016, the 26th Int. Conference on Computational Linguistics: Technical Papers*, 2016, pp. 1071–1082.

[60] M. Wang *et al.*, "A novel split-merge-evolve k clustering algorithm," in *2018 IEEE 4th Int. Conference on Big Data Computing Service and Applications (BigDataService)*. IEEE, 2018, pp. 229–236. [Online]. Available: https://doi.org/10.1109/BigDataService.2018.00041

[61] T. Mitchell *et al.*, "Never-ending learning," *Communications of the ACM*, vol. 61, no. 5, pp. 103–115, 2018. [Online]. Available: https://doi.org/10.1145/3191513

[62] I. Stoica *et al.*, "A berkeley view of systems challenges for ai," *arXiv preprint arXiv:1712.05855*, 2017.

[63] A. Tegen *et al.*, "Towards a taxonomy of interactive continual and multimodal learning for the internet of things," in *Adjunct Proceedings of the 2019 ACM Int. Joint Conference on Pervasive and Ubiquitous Computing and Int. Symposium on Wearable Computers*, 2019, pp. 524–528. [Online]. Available: https://doi.org/10.1145/3341162.3345603

[64] J. Konečnỳ *et al.*, "Federated learning: Strategies for improving communication efficiency," 2018.

[65] B. McMahan *et al.*, "Communication-efficient learning of deep networks from decentralized data," in *AI and statistics*. PMLR, 2017, pp. 1273–1282.

[66] T. Li *et al.*, "Federated learning: Challenges, methods, and future directions," *IEEE Signal Processing Magazine*, vol. 37, pp. 50–60, 2020. [Online]. Available: https://doi.org/10.1109/MSP.2020.2975749

[67] W.-T. Chang and R. Tandon, "Communication efficient federated learning over multiple access channels," *arXiv preprint arXiv:2001.08737*, 2020.

[68] E. Diao *et al.*, "Heterofl: Computation and communication efficient federated learning for heterogeneous clients," in *Int. Conference on Learning Representations*, 2021.

[69] A. Reisizadeh *et al.*, "Fedpaq: A communication-efficient federated learning method with periodic averaging and quantization," in *Int. Conference on AI and Statistics*. PMLR, 2020, pp. 2021–2031.

[70] Y. Chen *et al.*, "Communication-efficient federated deep learning with layerwise asynchronous model update and temporally weighted aggregation," *IEEE transactions on neural networks and learning systems*, vol. 31, no. 10, pp. 4229–4238, 2019. [Online]. Available: https://doi.org/10.1109/TNNLS.2019.2953131

[71] S. Caldas *et al.*, "Expanding the reach of federated learning by reducing client resource requirements," 2019. [Online]. Available: https://openreview.net/forum?id=SJlpM3RqKQ

[72] Y. Lin *et al.*, "Deep gradient compression: Reducing the communication bandwidth for distributed training," *Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017)*, 2017.

[73] F. Sattler *et al.*, "Robust and communication-efficient federated learning from non-iid data," *IEEE transactions on neural networks and learning systems*, vol. 31, no. 9, pp. 3400–3413, 2019. [Online]. Available: https://doi.org/10.1109/TNNLS.2019.2944481

[74] X. Wu *et al.*, "Fedmed: A federated learning framework for language modeling," *Sensors*, vol. 20, no. 14, p. 4048, 2020. [Online]. Available: https://doi.org/10.3390/s20144048

[75] A. Malekijoo *et al.*, "Fedzip: A compression framework for communication-efficient federated learning," *arXiv preprint arXiv:2102.01593*, 2021.

[76] D. Rothchild *et al.*, "Fetchsgd: Communication-efficient federated learning with sketching," in *Int. Conference on Machine Learning*. PMLR, 2020, pp. 8253–8265.

[77] J. Xu *et al.*, "Ternary compression for communication-efficient federated learning," *IEEE Transactions on Neural Networks and Learning Systems*, 2020. [Online]. Available: https://doi.org/10.1109/TNNLS.2020.3041185

[78] M. Asad *et al.*, "Ceep-fl: A comprehensive approach for communication efficiency and enhanced privacy in federated learning," *Applied Soft Computing*, vol. 104, p. 107235, 2021. [Online]. Available: https://doi.org/10.1016/j.asoc.2021.107235

[79] W. Luping *et al.*, "Cmfl: Mitigating communication overhead for federated learning," in *IEEE 39th international conference on distributed computing systems (ICDCS)*, 2019, pp. 954–964. [Online]. Available: https://doi.org/10.1109/ICDCS.2019.00099

[80] T. Nishio and R. Yonetani, "Client selection for federated learning with heterogeneous resources in mobile edge," in *IEEE ICC*. IEEE, 2019, pp. 1–7. [Online]. Available: https://doi.org/10.1109/ICC.2019.8761315

[81] S. Park *et al.*, "Fedpso: federated learning using particle swarm optimization to reduce communication costs," *Sensors*, vol. 21, no. 2, p. 600, 2021. [Online]. Available: https://doi.org/10.3390/s21020600

[82] Q. Xia *et al.*, "A survey of federated learning for edge computing: Research problems and solutions," *High-Confidence Computing*, vol. 1, no. 1, p. 100008, 2021. [Online]. Available: https://doi.org/10.1016/j.hcc.2021.100008

[83] A. A. Al-Saedi, V. Boeva, and E. Casalicchio, "Reducing communication overhead of federated learning through clustering analysis," in *2021 IEEE Symposium on Computers and Communications (ISCC)*. IEEE, 2021, pp. 1–7. [Online]. Available: https://doi.org/10.1109/ISCC53001.2021.9631391

[84] A. A. Al-Saedi, E. Casalicchio, and V. Boeva, "An energy-aware multi-criteria federated learning model for edge computing," in *2021 8th Int. Conf. on Future IoT and Cloud (FiCloud)*. IEEE, 2021, pp. 134–143. [Online]. Available: https://doi.org/10.1109/FiCloud49777.2021.00027

[85] D. L. Iverson, "Inductive system health monitoring," in *IC-AI*, 2004, pp. 605–611.

3^{rd} International Forum on Cyber Security, Privacy and Trust

NOWADAYS, information security works as a backbone for protecting both user data and electronic transactions. Protecting communications and data infrastructures of an increasingly inter-connected world have become vital nowadays. Security has emerged as an important scientific discipline whose many multifaceted complexities deserve the attention and synergy of computer science, engineering, and information systems communities. Information security has some well-founded technical research directions which encompass access level (user authentication and authorization), protocol security, software security, and data cryptography. Moreover, some other emerging topics related to organizational security aspects have appeared beyond the long-standing research directions.

The International Forum of Cyber Security, Privacy, and Trust (NEMESIS'22) as a successor of International Conference on Cyber Security, Privacy, and Trust (INSERT'19) focuses on the diversity of the cyber information security developments and deployments in order to highlight the most recent challenges and report the most recent researches. The session is an umbrella for all cyber security technical aspects, user privacy techniques, and trust. In addition, it goes beyond the technicalities and covers some emerging topics like social and organizational security research directions. NEMESIS'22 serves as a forum of presentation of theoretical, applied research papers, case studies, implementation experiences as well as work-in-progress results in cyber security. NEMESIS'22 is intended to attract researchers and practitioners from academia and industry and provides an international discussion forum in order to share their experiences and their ideas concerning emerging aspects in information security met in different application domains. This opens doors for highlighting unknown research directions and tackling modern research challenges. The objectives of the NEMESIS'22 can be summarized as follows:

- To review and conclude research findings in cyber security and other security domains, focused on the protection of different kinds of assets and processes, and to identify approaches that may be useful in the application domains of information security.
- To find synergy between different approaches, allowing elaborating integrated security solutions, e.g. integrate different risk-based management systems.
- To exchange security-related knowledge and experience between experts to improve existing methods and tools and adopt them to new application areas

Topics

- Biometric technologies
- Cryptography and cryptanalysis
- Critical infrastructure protection
- Security of wireless sensor networks
- Hardware-oriented information security
- Organization- related information security
- Social engineering and human aspects in cyber security
- Individuals identification and privacy protection methods
- Pedagogical approaches for information security education
- Information security and business continuity management
- Tools supporting security management and development
- Decision support systems for information security
- Trust in emerging technologies and applications
- Digital right management and data protection
- Threats and countermeasures for cybercrimes
- Ethical challenges in user privacy and trust
- Cyber and physical security infrastructures
- Risk assessment and management
- Steganography and watermarking
- Digital forensics and crime science
- Security knowledge management
- Security of cyber-physical systems
- Privacy enhancing technologies
- Trust and reputation models
- Misuse and intrusion detection
- Data hide and watermarking
- Cloud and big data security
- Computer network security
- Assurance methods
- Security statistics

Technical Session Chairs

- **Awad, Ali Ismail,** Luleå University of Technology, Sweden
- **Bialas, Andrzej,** Research Network Lukasiewicz – Institute of Innovative Technologies EMAG, Poland

DOI: 10.15439/2022F273
ISSN 2300-5963 ACSIS, Vol. 31

Rethinking Safety in Autonomous Ecosystems

David Halasz and Barbora Buhnova
Faculty of Informatics, Masaryk University
Brno, Czech Republic
{halasz, buhnova}@mail.muni.cz

***Abstract*—As autonomous cyber-physical systems are responding to the dynamism of our hyper-connected digital world, they are forming so called dynamic autonomous ecosystems, which require a change in methods ensuring their safe behavior. Within this change, reactions to predictable scenarios need to be replaced with adaptability to the unpredictable context, with gradual safety mechanisms, able to decide whether or not to trigger a certain mitigation procedure. In this paper, we outline our vision towards evolution of safety mechanisms to support dynamic and self-adaptive architectures of autonomous ecosystems. We are proposing an approach to address this research problem with the help of trust and reputation combined with gradual adaptation of safety procedures at runtime.**

I. Introduction

The growing demand for complex autonomous cyber-physical systems is stimulating their advancement in the direction of forming cooperative and collaborative autonomous ecosystems [1]. Such autonomous ecosystems, i.e. dynamic autonomous systems of systems, can provide a higher degree of autonomy and are capable of adapting to previously unknown situations. At the same time, however, their dynamic and self-adaptive nature is making it very challenging to ensure their safe and secure behavior, both at the individual level as well as at the level of the ecosystem as a whole [2].

The recent rapid development in autonomous driving is indicating that new autonomous systems might be joining city ecosystems sooner than the cities need to get ready to ensure the safety of these ecosystems as a whole, which is challenging not only technically but also from the perspective of understanding the ways in which societies perceive safety and trust in these autonomous systems.

Even though there has been substantial progress in the research of the methods ensuring safety in individual autonomous systems, the methods are falling short on the larger scale of autonomous ecosystems, in which the individual autonomous systems dynamically join and leave the ecosystem and interact with each other in a decentralized manner [1]. In this environment where multiple autonomous systems operate in the same physical space with high level of complexity and dynamic context changes, existing safety-assurance methods on the level of each individual systems might become too rigid to support the overall ecosystems.

Borrowing from the ways in which our societies ensure safety of its members, one of the most promising ways towards the safety of dynamic autonomous ecosystems is through the mechanisms of adaptive safety reflecting the actual safety risks in a given situation, which can be understood based on the trust and trustworthiness of the ecosystem members one interacts with [3], [2].

To stimulate the progress in this emerging field, the aim of this paper is to examine the problem of safety-assurance in dynamic autonomous ecosystems and envision an approach for adaptive safety in the ecosystems. Namely, we set the foundations for a new approach to adaptive safety, responding to the level of trust among autonomous systems. To this end, we first identify and present five scenarios of the key challenges related to safety in dynamic autonomous ecosystems, and then propose a framework to support adaptive safety in the ecosystems.

The structure of the paper as follows: Section 2 identifies the key challenges of adaptive safety in dynamic autonomous ecosystems and presents the example scenarios. Section 3 discusses research related to the addressed problem, followed with Section 4 presenting a solution and proposing a framework to support adaptive safety in dynamic autonomous ecosystems. Section 5 discusses assumptions and limitations made in designing the approach, which is followed with a conclusion and summary of future work.

II. Problem Description

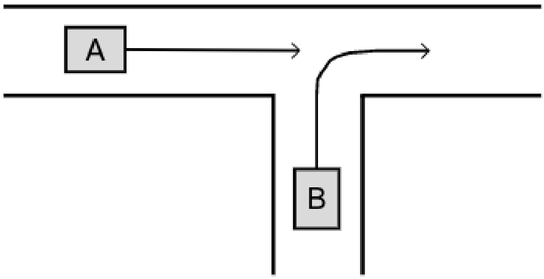
Safety as it is perceived and enforced on the level of individual autonomous cyber-physical systems is falling short on the magnitude of ecosystems [2]. The techniques that are capable of keeping a single system safe are not scaling to a dynamic ecosystem where member systems are heterogeneous and can join or leave the ecosystem at any time. This context requires an adaptive approach to safety that should be based on a kind of classification among member systems. One of the most promising strategies that only started to emerge recently, is to adapt safety to the level of trust among components within the whole ecosystem. In that context, a component of the autonomous ecosystem that is reported as untrusted by other ecosystem members (impacting its reputation within the ecosystem) might fall under (temporary) safety supervision and control, safeguarding its trustworthy operation.

To set the context for this research, this section identifies and discusses five challenges (described in the individual subsections) that need to be resolved to set the foundation of our approach to trust-driven adaptive safety in dynamic autonomous ecosystems. Our aim is to provide a solution to these challenges and propose a safety assurance framework for autonomous ecosystems on top of it.

A. Intentional vs. unintentional behavior

When an action done by an autonomous system has been classified as malicious, knowledge about the intent of this system can be an important decision factor when selecting the right kind of reaction. Unintentional malicious behavior can happen due to a malfunction in one of the components of the autonomous systems, delay in network communication or a software bug. On the other hand, it is possible to design a system that behaves in a harmful way in certain situations and tries to inflict as much damage as possible [2].

Fig. 1. Example scenario where intent classification is important



Consider an example in Figure 1 where two autonomous vehicles are meeting at an intersection. If vehicle B sends false information about its speed to vehicle A, relying on this information would cause a crash. In case the sensor responsible for measuring the speed of vehicle B is malfunctioning, a good course of action for vehicle A is to slow down and let vehicle B merge into the lane without interfering with it. However, if vehicle B is programmed to crash into vehicle A, the mitigation would not be sufficient in avoiding a crash.

B. Supervision awareness

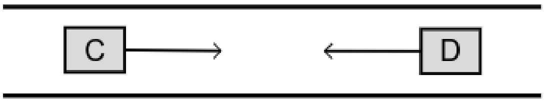
Unintentional malicious behavior can be sometimes corrected by letting the system know that it is behaving the wrong way. However, in some cases this operation would equip the system with additional information that could leveraged against another systems or to compromise the integrity of the ecosystem as a whole [2].

In the past decade, there were multiple scandals where vehicle manufacturers installed supervision awareness detection in their products [4]. The goal of this piece of software was to detect whether the vehicle is under emissions test or used by its owner. Based on the detected context, the ECU was instructed to reduce the CO_2 emissions by lowering the overall torque and power produced by the engine. This example can be simply extended to the domain to the autonomous ecosystems, where a member system can behave differently if it is being monitored and start behaving maliciously when it detects that the supervision is suspended.

C. Misclassification of behavior

When deciding whether an action of a system is malicious or not, there is always a margin of error. No classification technique can be always perfect and this inherently carries some danger when using such technique to make decision about enforcing safety mechanisms. If a behavior is incorrectly classified as malicious, reactions to this *false-negative* scenario can unnecessarily limit the functionality of the system. Furthermore, the result can influence any future interaction with this system can be restricted in the ecosystem. On the other hand, when a malicious action is wrongly classified as regular or safe behavior, this case is a *false-positive* and it can allow the system to cause even more damage then it originally intended to inflict on the ecosystem [2]. Some kind of compensation between these two extremes is necessary to both maintain the functionality and also ensure safety.

Fig. 2. Example scenario where misclassification can cause issues



Consider a scenario in Figure 2 where both misclassification cases can cause issues. If autonomous vehicle C is not capable of detecting the malicious intent of autonomous vehicle D, the situation can escalate into a frontal crash. Meanwhile, if vehicle D is wrongly classified as malicious, the triggered collision avoidance mechanism can slow down vehicle C more than it would be necessary with a correct classification. This would slow down the whole intersection for a longer amount of time that could affect other autonomous vehicles as well.

D. Feedback loops

The possibility of repetitive misclassification in multiple systems that interact with each other can create an even more challenging issue. Safety mechanisms invoked by one system can be interpreted by other systems as malicious. Any reaction to this false-negative can be also interpreted as malicious which can cause a gradual triggering of more and more strict safety features in every interacting system. This can even lead to a permanent stall of the whole ecosystem, especially if one of the member systems has been intentionally designed to cause an issue like this. This possibility should be considered when designing the safety architecture of an ecosystem [2].

Fig. 3. Example scenario where feedback loop can cause dangerous behavior of both vehicles

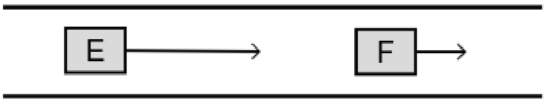
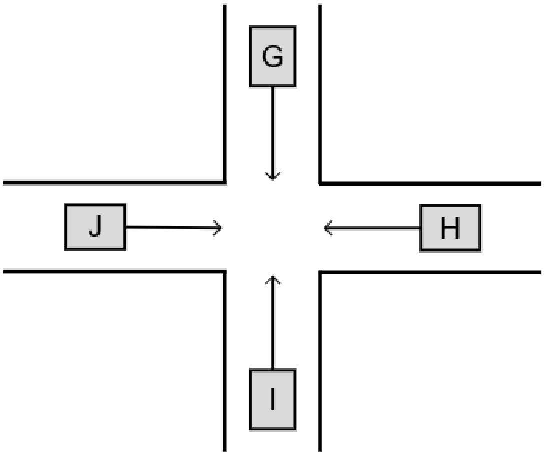


Figure 3 shows two autonomous vehicles heading in the same direction, where vehicle E has a higher speed rating and eventually it would overtake vehicle F. If vehicle F does not recognize the overtaking action and classifies the acceleration of vehicle E as malicious, it can accelerate to a higher speed to avoid a possible collision. This behavior can be interpreted by vehicle E as malicious and as a reaction it could decrease its speed. As there is no reason for vehicle F to increase its speed anymore, it can adjust its speed to the initial one. If vehicle E returns to its original speed, the whole situation can repeat itself from the beginning. The other possible outcome is to adjust the speed of both vehicles to the same speed, which would be not optimal for vehicle E as it is capable of higher speeds for a longer amount of time.

E. Compatibility

In any autonomous ecosystems there is a possibility of having heterogeneous member systems, manufactured by various vendors, using different implementations that not necessarily provide the same (safety) features. In order to ensure the safe behavior of the ecosystem, it is necessary to be able to provide some kind of backwards compatibility for systems with a reduced set of safety features. Alternatively, the ecosystem should be able to (at least temporarily) equip these systems with some kind of common safety mechanism. An extreme edge case of this problem is when a human is interacting with the ecosystem, which can be interpreted as a member system with zero compatibility and no possibility to receive a new safety mechanism.

Fig. 4. Example scenario for a critical compatibility issue



The example in Figure 4 shows four vehicles meeting at an intersection. Autonomous vehicles G and H use the most modern safety assurance framework which provides a solution for all the problems stated in Section 2. Autonomous vehicle I uses a different implementation in which some of these problems are not fully covered. Lastly, vehicle J is driven by a human who has no or minimal knowledge about what kind of software is running on the three autonomous vehicles.

III. State of the Art and Related Work

Safety can be interpreted differently in each domain. Our research found that the most relevant definitions for our purposes are "the ability of a distributed application and its parts to continue operating in a safe manner during and after a transformation" [5] and the "avoidance of hazards to the physical environment" [6].

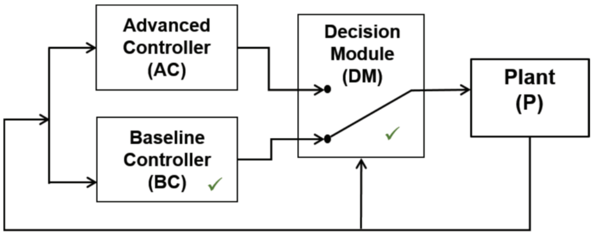
A. Safety in Autonomous Vehicles

Research in the area already covers most of the safety aspects of individual autonomous systems. Collision avoidance [7], communication security and recovering from attacks [8], [9] in the subdomain of autonomous vehicles are not dealing with safety on the magnitude of an ecosystem as a whole. Safety assurance in vehicle platooning [10] on the other hand is close to the area of interest, however, it does not provide answers to all the problems stated in Section 2.

B. Simplex architecture

The concept of "using simplicity to control complexity" [11] implemented by the Simplex architecture is an interesting approach to ensuring safe behavior of a system, popular in control systems and beyond. The core of the idea is to split up a system into a complex component (advanced controller) supporting all its ordinary behavior and a simpler component (baseline controller) that is only intended to resolve critical situations. A decision module between these two components can select which one should be enabled in certain situations [12]. While combining multiple simplexes can be a viable solution in having a complex system of systems where each system is responsible for its own safe behavior [13], [14], they are not designed to deal with uncertain situations. Also they can be prone to feedback loops and the lack of the granularity of the safety assurance can cause problems if a misclassification occurs [2].

Fig. 5. The simplex architecture [14]

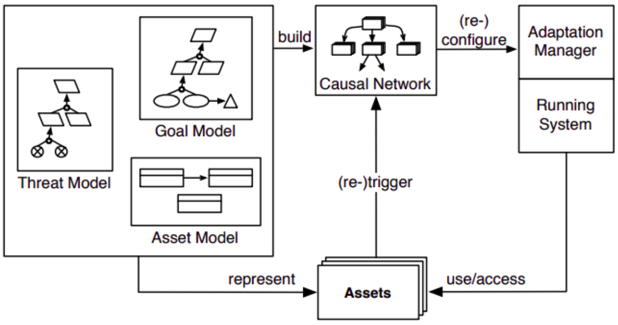


C. Self-adaptation

Self-adaptive cyber-physical systems are capable of handling uncertain situations [15]. This adaptability can be achieved by techniques such as runtime model querying [16] or Monitor Analyze Plan Execute with Knowledge (MAPE-K) [17] feedback loops. *Security* that is defined as "something "concerned with protecting assets from harm" can be also enforced in an adaptive way that is being evaluated during runtime [18]. This and techniques like Adaptive Control Lyapunov Functions aCLFs [19] can be also applied to (instead of security) enforce the *safety* of an autonomous system.

Although these solutions are well designed for autonomous systems, they do not provide answers to securing an ecosystem as a whole. Scaling a feedback loop by distributing it to multiple systems face difficulties as member systems are not always collaborative.

Fig. 6. Framework to support adaptive security [18]



D. Safety in distributed ecosystems

When considering safety in largely distributed ecosystems, wireless networks become a noteworthy source of knowledge, even though their most important aspect is communication security [20], [21]. The way how ad-hoc and self organizing mesh networks [22], [23] work strongly resembles the dynamicity in autonomous ecosystems. Techniques from this subdomain [24], [25], [26] might be useful in our context. It is however limiting that they can only cover the cyber part of a cyber-physical ecosystem. Cutting off a node from a network can surely increase safety, however, ignoring a robot or an autonomous vehicle in a similar way can cause more problems than it solves. Moreover, such scenario could hint a malicious system about no longer being monitored, which introduces the problem of supervision awareness in the ecosystem.

E. Classification of behavior

Determining if an action done by a system is malicious or not is the most important factor when selecting the appropriate reaction in other member systems. Due to the dynamicity of the ecosystem, such classification has to be conducted continuously and at real-time. In an ideal case, each member system could have its own model constructed [27] and propagated that could be queried by other systems to decide on further actions. The approach is called *models@run.time* [16] and it is intended to be used in scenarios that were not taken into account when the system had been designed [28]. The problem with this approach is the requirement of a valid model for each member system, which is not always possible to construct. Another issue is the distribution of these models and the fact that sharing them with malicious systems might equip them with knowledge about vulnerabilities and increase the overall attack surface [2].

IV. Proposed solution

Ensuring safe and secure behavior of autonomous cyber-physical ecosystems is a challenging task and it requires a new approach on how relationships between individual entities are perceived. The inherent complexity and the dynamic context changes of the consistency of such ecosystems cannot be solved during design time. Therefore, any proposed solution has to able to handle previously unexpected or uncertain situations during runtime. Referring to our previous work [2] and Liu et al. [29], we believe that leveraging real-time evaluated trust among ecosystem components can provide sufficient input to make real-time decisions about safety in dynamic autonomous ecosystems [30].

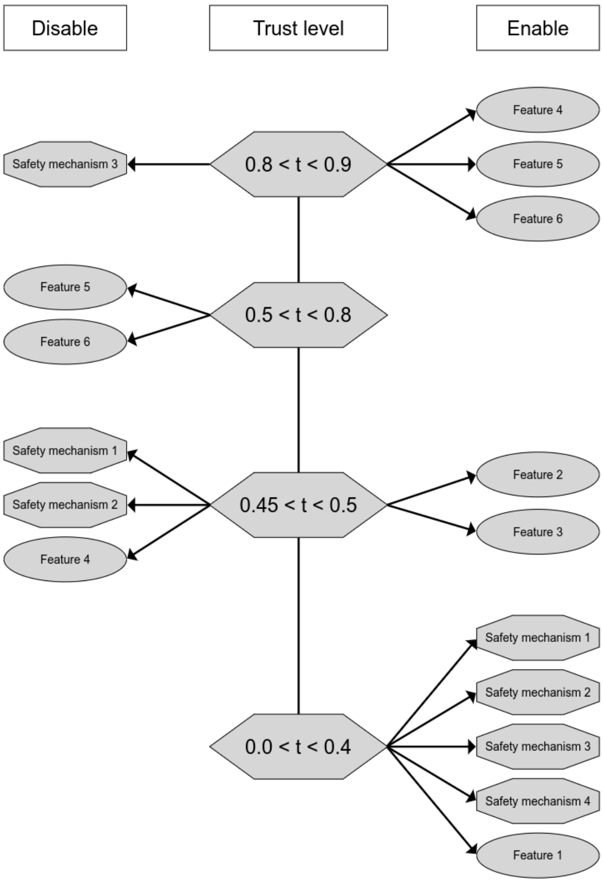
The definition of trust can to be borrowed from different branches of science [31], such as Psychology [32], Philosophy [33] and Organizational Management [34]. Simply put, autonomous systems shall understand trust similarly as we humans do.

Most of the research in the area is conducted around the qualitative understanding of trust [31], i.e. classifying it into a binary form to either trust or not to trust. These approaches, however, due to the lack of the granularity of their output, are prone to misclassification. Since the appearance of the Internet of Things, there are some promising approaches that are able to assess trust quantitatively [29], e.g. into a percentage. Due to the higher variety in the output, such methods are statistically less likely to be far away from the right result in case of an error in the trust calculation. It is important to mention, that trust can be calculated directly (trust) from a target system or obtained indirectly (reputation) from other systems that have had former interactions with the target system [35], [36], [37]. In some cases the two can be merged into a combined value with predefined or dynamic weights.

Any input consumed by our solution has to be more granular than a binary, e.g. *to trust* or *not to trust*. This should significantly reduce the chance of errors happening due to misclassification as the distance between the ideal and the actual output is statistically lower than in a binary situation. This granularity should be mirrored in the safety enforcement with a graduality of triggering safety mechanisms or exposing features towards other autonomous systems [2].

The decision tree for a single autonomous system of an example safety mechanism is shown in Figure 7. For simplification, it assumes a numeric input coming from a trust model with a decimal number between 0 and 1. This trust level is being continuously recalculated in real-time against any interacting autonomous system and used to decide which features should be exposed or concealed and which safety mechanisms should be triggered in individual situations. A low trust level allows a minimal set of features and a large number of safety mechanisms and as the trust level is growing, the trend is gradually reversing. It is important, that safety mechanisms are also available on the highest level of trust and the system can move to a lower trust level at any time during its operation.

Fig. 7. Example decision tree: actions to take based on trust [2]



A. Safety Assurance Framework

The envisioned framework supporting trust-based adaptive safety is drafted in Figure 8. Its core components are the Trust Model, the Decision Tree and the Safety Module connecting these two. 1) The *Trust Model* calculates a *Trust Value* of a target system based on inputs from sensors and a reputations propagated by other actors of the ecosystem. 2) The *Trust Value* calculated by the model is consumed by the Safety Module and propagated to other systems doing similar calculations. 3) The *Safety Module* using the Decision Tree selects what safety mechanisms should be enabled or disabled and what features can be exposed to or concealed from the target system. 4) The *Decision Tree* allows the possibility to alter itself either by a software update or by the system itself using a self-adaptive technique.

The process of trust calculation is continuously triggered by the *Safety Module* after each adaptation cycle. This ensures that the system has the most recent information about how much a target system can be trusted at all time.

In case the *Trust Value* has been assessed wrongly, granularity of the trust output combined with the gradual triggering of safety features significantly decreases the margin of error and its consequences. In case a system with high level of trust behaves maliciously, the fast recalculation of the *Trust Value* can not just help the attacked system to quickly adapt, but this new information can be propagated to other members of the ecosystem. The spreading reputation can influence trust computations in other systems, ensuring their safety features are prepared for a future encounter with a malicious system. Furthermore, fresh reputation information might provide means to end a feedback loop introduced by wrong trust calculations.

The real-time recalculation and quick reaction time of triggering safety features should address most of the possible issues related to the detection of intent. Due to the continuous recalculation of the *Trust Value*, supervision awareness is also addressed by this technique. A malicious system that previously maximized its trustworthiness to access a certain feature can be quickly detected. We envision that trust propagation in ecosystems would create clusters of safely operating member systems and push malicious ones to the periphery. Interaction with them should not be completely severed as knowledge about them can be helpful in preventing further harm in the future.

If trust is calculated by using predictive simulations via Digital Twins [3], the same Digital Twins can be also used to partially determine the capabilities of other systems. This equips the Safety Assurance Framework with critical information regarding compatibility and might prevent certain cases of feedback loops. In this case the framework has to be able to receive Digital Twins and run predictive simulations even independently from its *Trust Model*. It is also necessary to consider situations when digital twins are not available or predictive simulation is not an option, e.g. in case of a human. Most importantly, the *Decision Tree* has to be constructed in a way that it handles these situations.

B. Example scenario

Consider a scenario of two autonomous vehicles in Figure 9 leveraging our framework. Both vehicles can in advance assess how much they trust each other. In case if that information is available, they can rely on reputation propagated by other actors of the ecosystem as well. From the perspective of vehicle A, when trust towards the malicious vehicle B is calculated as 0.7, the system would initially trigger safety mechanisms only for avoiding a frontal collision by moving to the right side of the road. As the *Trust Value* is high enough, this information would be communicated to vehicle B. If the malicious behavior is becoming more obvious, the constantly recalculated *Trust Value* first drops to 0.4, vehicle A begins to reduce its speed and tries to find a course that would minimize the risk of a collision. As the vehicles are getting closer to each other and the *Trust Value* reaches 0.2, an active collision avoidance mechanism takes over the control and tries to keep safe distance from vehicle B.

In a reversed scenario if the wrongly calculated *Trust Value* is 0.2, the same active collision avoidance mechanism is controlling vehicle A. Vehicle B detects this behavior and

Fig. 8. Framework to support Trust-based Adaptive Safety

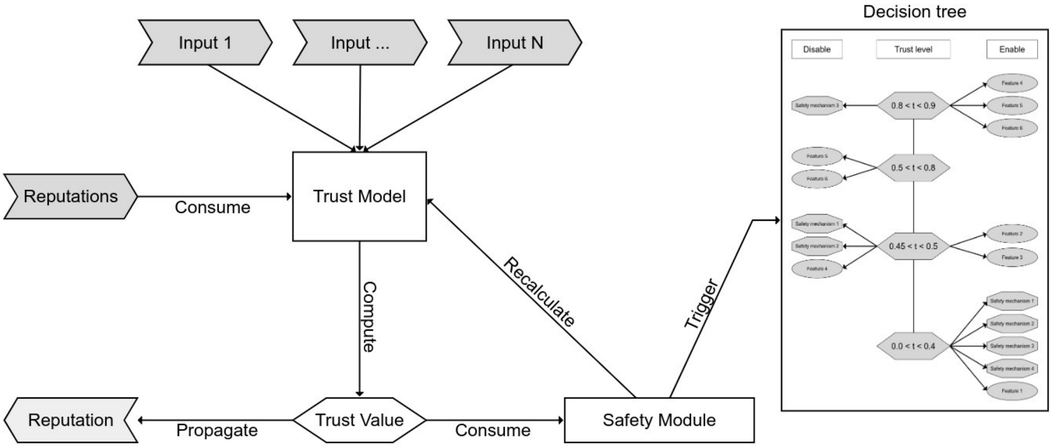
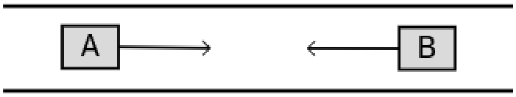


Fig. 9. Example scenario



also starts behaving more cautiously, which increases the *Trust Value* towards it to 0.5. This leads to more information sharing between the two vehicles and the increased number of inputs increases the *Trust Value* to 0.9 by the time the two vehicles meet and they both move to their right side of the road and continue with an increased speed.

After their encounter, both vehicles store the final *Trust Value*, that would prevent them from getting into similar scenarios with each other. In the meantime, all calculated *Trust Values* are propagated into the ecosystem (in terms of a reputation of each individual vehicle), which should reduce any misclassification for other actors.

V. Discussion

Our approach proposes a paradigm shift in comparison with existing solutions. This paper is exploratory in nature and intends to start a community discussion about future steps in this direction.

Even though trust is the designated decision factor in our approach, it is only considered as an input to the proposed mechanism. The *Trust Model* is treated as a black box and its main requirement is to produce a non-binary granular output. Having a safety mechanism decoupled from its input allows us to have additional flexibility. Autonomous systems can implement different *Trust Models* [30] and it might also happen that decision factors other than trust will be consumed by the solution.

Our future plans are to finalize the specification of the Safety Assurance Framework and define its input and output interfaces. Meanwhile, our research team is reviewing trust computation methods that can be consumed by the proposed solution. After both are ready and available, our plan is to reach out to automotive companies and work together with them to validate the framework on real-life case studies.

VI. Conclusion

Due to the rising complexity of autonomous ecosystems, new software-architecture mechanisms are necessary to respond to the dynamicity of changes while ensuring safety even in uncertain situations. In this work, we propose to address this challenge via mechanisms to gradually enable safety mechanisms based on the assessed level of trust towards other members of the ecosystem, in combination to new ways of assessing the trustworthiness of individual system components. Furthermore, we describe the fundamentals of a Safety Assurance Framework that would support this mechanism. In our next steps we plan to create a more thorough design of the framework and validate it on case studies provided by possible partners from the industry. We believe that this idea will evolve into a set of prototypical tools supporting promoting of safe autonomous ecosystems.

Acknowledgment

This research was supported by ERDF "CyberSecurity, CyberCrime and Critical Information Infrastructures Center of Excellence" (No. CZ.02.1.01/0.0/0.0/16_019/0000822).

References

[1] R. Capilla, E. Cioroaica, B. Buhnova, and J. Bosch, “On autonomous dynamic software ecosystems,” *IEEE Transactions on Engineering Management*, pp. 1–15, 2021. doi: 10.1109/TEM.2021.3116873

[2] D. Halasz, “From systems to ecosystems: Rethinking adaptive safety,” in *17th International Symposium on Software Engineering for Adaptive and Self-Managing Systems (SEAMS ’22).* IEEE, 2022. doi: 10.1145/3524844.3528067

[3] E. Cioroaica, T. Kuhn, and B. Buhnova, “(do not) trust in ecosystems,” in *2019 IEEE/ACM 41st International Conference on Software Engineering: New Ideas and Emerging Results (ICSE-NIER).* IEEE, 2019. doi: 10.1109/ICSE-NIER.2019.00011 pp. 9–12.

[4] M. Contag, G. Li, A. Pawlowski, F. Domke, K. Levchenko, T. Holz, and S. Savage, "How they did it: An analysis of emission defeat devices in modern automobiles," in *2017 IEEE Symposium on Security and Privacy (SP)*, 2017. doi: 10.1109/SP.2017.66 pp. 231–250.

[5] P. McKinley, S. Sadjadi, E. Kasten, and B. Cheng, "Composing adaptive software," *Computer*, vol. 37, no. 7, pp. 56–64, 2004. doi: 10.1109/MC.2004.48

[6] A. Banerjee, K. K. Venkatasubramanian, T. Mukherjee, and S. K. S. Gupta, "Ensuring safety, security, and sustainability of mission-critical cyber–physical systems," *Proceedings of the IEEE*, vol. 100, no. 1, pp. 283–299, 2012. doi: 10.1109/JPROC.2011.2165689

[7] G. Li, Y. Yang, T. Zhang, X. Qu, D. Cao, B. Cheng, and K. Li, "Risk assessment based collision avoidance decision-making for autonomous vehicles in multi-scenarios," *Transportation Research Part C: Emerging Technologies*, vol. 122, p. 102820, 2021. doi: 10.1016/j.trc.2020.102820. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0968090X20307257

[8] S. Bouchelaghem, A. Bouabdallah, and M. Omar, "Autonomous Vehicle Security: Literature Review of Real Attack Experiments," in *The 15th International Conference on Risks and Security of Internet and Systems*, Paris, France, 2020. [Online]. Available: https://hal.archives-ouvertes.fr/hal-03034640

[9] J. Cui, L. S. Liew, G. Sabaliauskaite, and F. Zhou, "A review on safety failures, security attacks, and available countermeasures for autonomous vehicles," *Ad Hoc Networks*, vol. 90, p. 101823, 2019. doi: 10.1016/j.adhoc.2018.12.006 Recent advances on security and privacy in Intelligent Transportation Systems. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1570870518309260

[10] J. Axelsson, "Safety in vehicle platooning: A systematic literature review," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, pp. 1–13, 08 2016. doi: 10.1109/TITS.2016.2598873

[11] L. Sha, "Using simplicity to control complexity," *IEEE Software*, vol. 18, no. 4, pp. 20–28, 2001. doi: 10.1109/MS.2001.936213

[12] D. Seto, B. Krogh, L. Sha, and A. Chutinan, "The simplex architecture for safe online control system upgrades," in *Proceedings of the 1998 American Control Conference. ACC (IEEE Cat. No.98CH36207)*, vol. 6, 1998. doi: 10.1109/ACC.1998.703255 pp. 3504–3508 vol.6.

[13] P. Vivekanandan, G. Garcia, H. Yun, and S. Keshmiri, "A simplex architecture for intelligent and safe unmanned aerial vehicles," in *2016 IEEE 22nd International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA)*, 2016. doi: 10.1109/RTCSA.2016.17 pp. 69–75.

[14] D. Phan, J. Yang, M. Clark, R. Grosu, J. Schierman, S. Smolka, and S. Stoller, "A component-based simplex architecture for high-assurance cyber-physical systems," in *2017 17th International Conference on Application of Concurrency to System Design (ACSD)*, 2017. doi: 10.1109/ACSD.2017.23 pp. 49–58.

[15] H. Muccini, M. Sharaf, and D. Weyns, "Self-adaptation for cyber-physical systems: A systematic literature review," in *2016 IEEE/ACM 11th International Symposium on Software Engineering for Adaptive and Self-Managing Systems (SEAMS)*, 2016. doi: 10.1145/2897053.2897069 pp. 75–81.

[16] N. Bencomo, R. France, B. Cheng, and U. Aßmann, Eds., *Models@run.time: foundations, applications, and roadmaps*, ser. Lecture notes in computer science. Germany: Springer, Dec. 2014. ISBN 978-3-319-08914-0 Dagstuhl Seminar 11481 on models@run.time held in November/December 2011.

[17] P. Arcaini, E. Riccobene, and P. Scandurra, "Modeling and analyzing mape-k feedback loops for self-adaptation," in *2015 IEEE/ACM 10th International Symposium on Software Engineering for Adaptive and Self-Managing Systems*, 2015. doi: 10.1109/SEAMS.2015.10 pp. 13–23.

[18] M. Salehie, L. Pasquale, I. Omoronyia, R. Ali, and B. Nuseibeh, "Requirements-driven adaptive security: Protecting variable assets at runtime," in *2012 20th IEEE International Requirements Engineering Conference (RE)*, 2012. doi: 10.1109/RE.2012.6345794 pp. 111–120.

[19] A. J. Taylor and A. D. Ames, "Adaptive safety with control barrier functions," in *2020 American Control Conference (ACC)*, 2020. doi: 10.23919/ACC45564.2020.9147463 pp. 1399–1405.

[20] L. S. Rutledge and L. J. Hoffman, "A survey of issues in computer network security," *Computers & Security*, vol. 5, no. 4, pp. 296–308, 1986. doi: 10.1016/0167-4048(86)90050-7. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0167404886900507

[21] M. S. Siddiqui, "Security issues in wireless mesh networks," in *2007 International Conference on Multimedia and Ubiquitous Engineering (MUE'07)*, 2007. doi: 10.1109/MUE.2007.187 pp. 717–722.

[22] A. J. Fehske, I. Viering, J. Voigt, C. Sartori, S. Redana, and G. P. Fettweis, "Small-cell self-organizing wireless networks," *Proceedings of the IEEE*, vol. 102, no. 3, pp. 334–350, 2014. doi: 10.1109/JPROC.2014.2301595

[23] R. Katulski, J. Stefański, J. Sadowski, S. Ambroziak, and B. Miszewska, *Self-Organizing Wireless Monitoring System for Containers*. Springer, 08 2009, pp. 164–172. ISBN 978-3-642-03840-2

[24] S. Desilva and R. Boppana, "Mitigating malicious control packet floods in ad hoc networks," in *IEEE Wireless Communications and Networking Conference, 2005*, vol. 4, 2005. doi: 10.1109/WCNC.2005.1424844 pp. 2112–2117 Vol. 4.

[25] M.-Y. Su, K.-L. Chiang, and W.-C. Liao, "Mitigation of black-hole nodes in mobile ad hoc networks," in *International Symposium on Parallel and Distributed Processing with Applications*, 2010. doi: 10.1109/ISPA.2010.74 pp. 162–167.

[26] A. Naveena and K. R. L. Reddy, "Malicious node prevention and mitigation in manets using a hybrid security model," *Information Security Journal: A Global Perspective*, vol. 27, no. 2, pp. 92–101, 2018. doi: 10.1080/19393555.2017.1415399. [Online]. Available: 10.1080/19393555.2017.1415399

[27] S. Kent, "Model driven engineering," in *Integrated Formal Methods*, M. Butler, L. Petre, and K. Sere, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002. ISBN 978-3-540-47884-3 pp. 286–298.

[28] M. Barkowsky, T. Brand, and H. Giese, "Improving adaptive monitoring with incremental runtime model queries," in *2021 International Symposium on Software Engineering for Adaptive and Self-Managing Systems (SEAMS)*, 2021. doi: 10.1109/SEAMS51251.2021.00019 pp. 71–77.

[29] L. Liu, M. Loper, Y. Ozkaya, A. Yasar, and E. Yigitoglu, "Machine to machine trust in the iot era," in *Proceedings of the 18th International Conference on Trust in Agent Societies - Volume 1578*, ser. TRUST'16. Aachen, DEU: CEUR-WS.org, 2016, p. 18–29.

[30] D. Iqbal and B. Buhnova, "Model-based approach for building trust in autonomous drones through digital twins," in *2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2022.

[31] J.-H. Cho, K. Chan, and S. Adali, "A survey on trust modeling," *ACM Comput. Surv.*, vol. 48, no. 2, oct 2015. doi: 10.1145/2815595. [Online]. Available: 10.1145/2815595

[32] J. B. Rotter, "Interpersonal trust, trustworthiness, and gullibility." *American psychologist*, vol. 35, no. 1, p. 1, 1980. doi: 10.1037/0003-066X.35.1.1

[33] B. Lahno, "On the emotional character of trust," *Ethical theory and moral practice*, vol. 4, no. 2, pp. 171–189, 2001.

[34] R. C. Mayer, J. H. Davis, and F. D. Schoorman, "An integrative model of organizational trust," *The Academy of Management Review*, vol. 20, no. 3, pp. 709–734, 1995. doi: 10.2307/258792. [Online]. Available: http://www.jstor.org/stable/258792

[35] A. I. A. Ahmed, S. H. Ab Hamid, A. Gani, M. K. Khan *et al.*, "Trust and reputation for internet of things: Fundamentals, taxonomy, and open research challenges," *Journal of Network and Computer Applications*, vol. 145, p. 102409, 2019. doi: 10.1016/j.jnca.2019.102409

[36] A. Hbaieb, S. Ayed, and L. Chaari, "A survey of trust management in the internet of vehicles," *Computer Networks*, vol. 203, p. 108558, 2022. doi: 10.1016/j.comnet.2021.108558. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1389128621004758

[37] W. Abdelghani, C. A. Zayani, I. Amous, and F. Sèdes, "Trust management in social internet of things: a survey," in *Conference on e-Business, e-Services and e-Society*. Springer, 2016. doi: 10.1007/978-3-319-45234-0_39 pp. 430–441.

DOI: 10.15439/2022F272
ISSN 2300-5963 ACSIS, Vol. 31

A Comparative Study of User Identification for COVID-19 Vaccination Online Registration

Sota Kato
Department of Informatics
The University of Electro-Communications
1-5-1 Choufugaoka, Chofu, Tokyo 182-8585, Japan
Email: k1910165@edu.cc.uec.ac.jp

Soontorn Sirapaisan, Chalee Vorakulpipat
Information Security Research Team
National Electronics and Computer Technology Center
Pathumthani, Thailand
Email: soontorn.sir@nectec.or.th, chalee.vor@nectec.or.th

***Abstract*—Due to the COVID-19 pandemic, individuals have been encouraged to obtain COVID-19 vaccinations. As part of the vaccination process, it is necessary to verify the identity before obtaining the vaccines. Incorrectly identifying the individuals could lead to serious consequences. For example, an incorrect vaccination schedule (e.g., incorrect timing between doses and incompatible brands) may cause undesired side effects to an individual, so it should be more accurate and convenient to identify themselves whenever and wherever people are required. One of the most effective and commonly used identification methods is face-to-face (or offline) identification. However, the method is typically time-consuming and is not suitable for the current situation where individuals should avoid direct contact. Hence, there is a growing trend for an online identification method where individuals use digital credentials to identify themselves. This paper suggests an online authentication system for improvements by investigating how different people authenticate themselves online in each country for COVID-19 vaccination. Comparing the online authentication systems between a country that have their national identification cards or not makes the Pros and Cons of both systems clear, particularly, in Japan and Thailand.**

***Index Terms*—COVID-19, Online Identification, Digital ID**

I. Introduction

SINCE the COVID-19 is spread, it has been recommended to have the vaccination in most countries. The vaccination process must be correct to identify each person and make it easy to vaccinate many kinds of people. Currently, the reservation process is mainly online to avoid many citizens from gathering in person. The notable issues are to distinguish whether those who have done the required number of vaccination or not clearly in the defined situation and control individual vaccine data, authenticating themselves. That is why what to certify must associate with a national database to deal with data efficiently in that situation.

The governments of each country are required urgent different responses based on their identification contexts. For example, most people in the USA use mainly Social Security Number(SSN) to identify themselves, and Japan has identity cards as well, however, some people do not have a way to identify. On the other hand, almost all people in Thailand have a unique national identity card. Therefore, Japan has different contrast contexts to Thailand in identifying each person because there is no obligation to have a unique national identity card like those in Thailand have.

The main contributions of this paper are three-fold. The first is an investigation of common identification methods commonly used to identify individuals. The second is a detailed comparison between the identification methods used in Japan and Thailand to understand their advantages and disadvantages and how the methods work. The third is the proposal of an online identification method that is more secure and as convenient to use as the methods currently used in Japan and Thailand.

The remaining of this paper is organized as follows. In Section II, we introduce related works about basic issues the current identity cards have. In Section III, we summarize general authentication system in Japan, also present the COVID-19 vaccination flow. In Section IV, we analyze authentication levels based on situations in Thailand and explain the COVID-19 vaccination flow. In Section V, the Pros and Cons of authentication systems are presented. In Section VI, we recommend an online authentication system for improvements. In Section VII, we conclude this paper and descrive further work.

II. Literature review

Traditionally, identity cards are used to authenticate people in most all countries[1]. In Japan, there are many kinds of identity cards. Some of them are easily falsified because counterfeiting businesses exist, so it can not be determined if they are authentic or not. The effective measurement is to complement personal information without violating human rights. Identity cards should not refer to just name and biometric information but also a person's background. Additionally, to prevail identity cards with integrated circuit chips results in preventing them from counterfeiting, however, they have not spread well yet. Discussing how methods to use authentication can sort out these problems and realize correct authentication, compared to Thailand's authentication methods.

The certification of residence is used as a major social infrastructure in Japan, which means people in Japan are controlled by each province depending on their residence[2]. Every city hall manages the data and uses them for authentication on each person when enforcing public service, however, the system could not work in Tohoku earthquake (11 March 2011) because of the power outage. Under this situation, it takes a lot of time to distinguish people who lose every identity card due to the disaster in addition to not generalizing a formal identity card. Volunteer medical workers confirm in person the patient's name and address[3]. Predictably, the current certification method will shift from physical certificates in visible form to media-information independent of them. A better way of authentication should work in these circumstances.

III. User identification in Japan

In general and formal, the number of required identity certifications is up to the difficulty to obtain and how accurate they are.

TABLE I
One is sufficient identification

One document	
· A driver's license	· A Passport
· A individual number card	· A welfare certification
· A certification of driving history	· A residence card for foreiner

There are a driver's license, a passport, and so on as one sufficient document to identify. Their specification is to attach a facial photo in addition to individual information.

TABLE II
Two are sufficient identification

Two documents	
mandatory one of them at least	any identify cards
· A certification of insurance	· A certification of residence
· A certification of national pension	· A issue from governments
· A mother and child health handbook	· A receipt of social insurance
· A certification of the stamp used for signature	· A receipt of public utility

On the other hand, if more than one document is needed for a person's authentication, at least one of them has to be from the list of mandatory documents, e.g. the certification of insurance or the certification of national pension. Japanese people have an obligation that they join one of the various insurance and pay a pension, so obtaining certifications like these is not harder than obtaining a driver's license and a passport. For example, when people create banking accounts on sites, in addition to those kinds of identification cards, people need a name stamp used for their signature and required identification to verify their name, address, and birthday. That is why can authenticate individuals more correctly because if they do not have the same name stamp used to enroll in the bank account, they can not authenticate themselves. They must inform the bank that they have lost their name stamp and create a new one in the name stamp shop. On the other hand, when people apply for a bank account online, they need two identity cards as follows. The required documents are more rigorous than the general situation.

- A driver's license
- A passport
- A individual number card
- A certification of insurance
- A residence card for a foreigner [1]

First, people take a picture of identity cards. Next, they fill out the blank regarding name, address, mail address, and place of work. Finally, SMS is sent by websites, then they confirm the numbers.

Hence, if people do not have any identity cards, they can not authenticate themselves in the required situation although it is not mandatory to have regularly their identity cards every time. Those who have a residence in Japan indeed have individual numbers for each person like the Thailand system, even if they are babies and foreigners, however, the individual numbers currently are required in only limited situations such as when you are hired. In conclusion, the main cards used as identification are the certification of insurance, the certification of residence, and a driver's license.

People must authenticate themselves in also COVID-19 vaccine situation. The incorrect certification caused some problems like the unexpected doses of different vaccine makers and abnormal time of vaccination, but high authentication security disturbs people from vaccinating conveniently. In Japan's vaccination process, the necessary factors are to have the certification of residence and a vaccine ticket written with individual numbers for the COVID-19 vaccine. When people vaccinate, they can even use only a certification of insurance without combining it with other identity documents, which becomes easier to vaccinate every people, unlike the general situation.

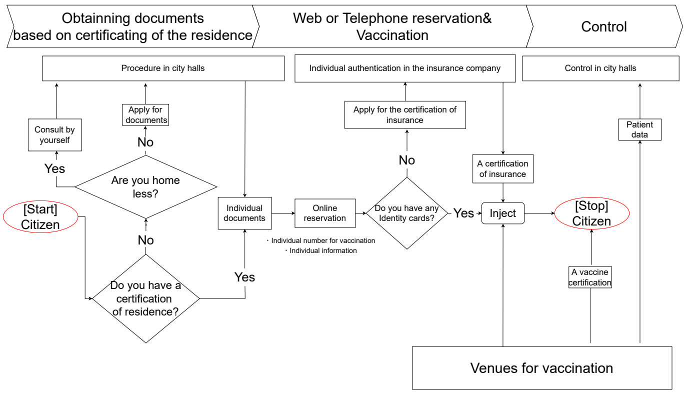


Fig. 1. Japanese vaccination flowchart

First, citizens could obtain the documents which include a vaccine ticket written with individual numbers for the COVID-19 vaccine, a medical history form, and guidance notifications from their city halls in their homes for the first time dose. That is why people who are homeless can not obtain them because they do not have the certification of residence, so they need to consult with city hall clerks by themselves. Only if people are not homeless, there are situations in that people could not obtain them automatically as follows.

- Japanese people who are not currently living in their home districts. This includes those that move to other

[1] It can not be combined with a passport

districts recently and those that are currently staying abroad.
- Foreigners who are staying in Japan for more than three months.

These people have to apply for documents to their city halls by themselves, and their city halls mail the documents to them.

After obtaining a vaccine ticket, there are mainly two ways of reserving a vaccination in Japan. One is that people could reserve the vaccination date and the venue through websites, filling the blanks requiring individual numbers for COVID-19 vaccine, name, birthday, phone number, and mail address. The other is through the telephone. People are asked about the same information as websites. In both ways, they do not have to show real identity cards, unlike online creating bank accounts. When they vaccinate in the hospital, they must bring the vaccine ticket, medical history form, and identity cards to identify themselves before vaccinating and controlling vaccine information. Exactly, all Japanese people have no less than the certification of insurance as identity cards because entering insurance is an obligation in Japan if only they are losing it or in the process of changing insurance. In any situation, they can obtain their one immediately based on insurance companies.

Finally, hospitals distribute and attach the vaccine certification to the vaccine ticket for people who have done vaccination, and it is used for authentication for second-time doses, combining individual numbers for COVID-19 and identity cards. Moreover, Hospitals have to inform city halls of patient data for COVID-19 vaccination. According to that, city halls control them and prevent from sending plural vaccine tickets to the same person. Since the end of last year, it has become possible to obtain vaccine certification through an app thanks to a digital agency organized by the Japanese government. To use it, people must have their individual number card with an integrated circuit chip, because the app use "near field communication (NFC)" technology through an integrated circuit chip. That is why people need to have NFC smartphones compatible with an individual number card.

IV. User identification in Thailand

In Thailand, every person who is at least 7 years old is legally obliged to carry a national identity card whenever outside. The national identity card is made from plastic with an integrated chip like a Japanese individual number card. While Japanese governments have tried to prevail against it, a similar card has already spread in Thailand. It is used for a variety of situations such as when people create bank accounts, obtain a driver's license, and enroll in insurance. The Thailand national card has included these information as follows, and which are controlled by Department of Provincial Administration(DOPA).

- Name
- Gender
- Birthday
- Religion
- Blood type
- Address
- 13 digits individual number

Every transaction is held online usually, using them in general. Electronic Transactions Development Agency(ETDA) has made a definition of what kinds of identification citizens should need based on situations. It is called Identity Assurance Level(IAL)[4]. The certification levels are different though people use mainly their national identity cards.

TABLE III
The IAL layers

IAL levels	What people need
IAL 1.1	No identification
IAL 1.2	A copy of the national identify card A copy of a passport
IAL 1.3	The real national identify card A real passport
IAL 2.1	The integrated chip data extracted from the national identify card and confirming phone numbers Reading the NFC data sent from a passport and a taking facial photo
IAL 2.2	The integrated chip data extracted from the national identify card, cheking authentication with DOPA data base and confirming phone numbers Reading the NFC data sent from a passport and taking a facial photo
IAL 2.3	The integrated chip data extracted from the national identify card,cheking authentication with DOPA data base and confirming phone numbers Reading the NFC data sent from a passport and taking a photo of IC chip and facial photo, comparing with biometric certification
IAL 3	The integrated chip data extracted from the national identify card,cheking authentication with DOPA data base, confirming phone numbers and meeting in person or virtual

Unlike Japan's identification system, whenever need to certify, citizens use the unique identification card in Thailand. To select the most appropriate identification[5], IAL levels help agencies decide which ways to identify people are appropriate for their digital service demands. Particularly, in the COVID-19 situation, the level of online reservation and vaccination protocols are included around IAL 2.

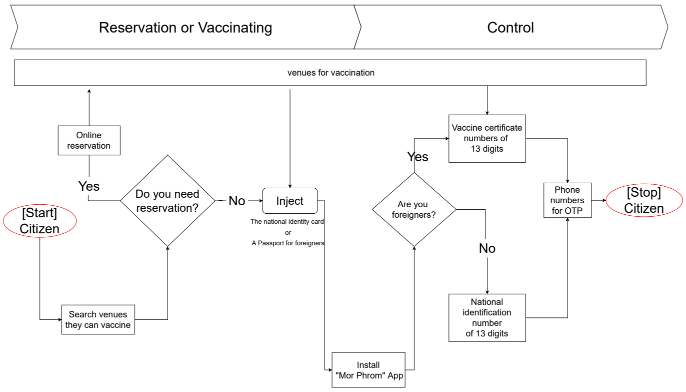

Reservation or Vaccinating
Control
venues for vaccination
Online reservation
Yes
Do you need reservation?
No
Inject
The national identity card or A Passport for foreigners
[Start] Citizen
Search venues they can vaccine
Install "Mor Phrom" App
Are you foreigners?
Yes
Vaccine certificate numbers of 13 digits
No
National identification number of 13 digits
Phone numbers for OTP
[Stop] Citizen

Fig. 2. Thailand vaccination flowchart

First, citizens search the venues for vaccination and could go there even if they do not reserve in advance as well as foreigners. Next, it is required to show only the national identity card or passport to identify in the venues. This is how they can vaccinate smoothly for the first time. Finally, "Mor Phrom" is spread to certify the COVID-19 vaccination. Citizens need only three steps with the required identity documents. The first step is to enter 13 digits numbers with the national identity card, so they are used instead of the individual numbers issued for a vaccination ticket in Japan. In the case of foreigners, the venues issue the specific 13 digits numbers for the vaccine certification. The next step is to enter the phone number to send One Time Password(OTP). After that, once set the original password, it certifies the vaccination information whenever and wherever citizens want to show it. Therefore, the level of the vaccine certification process is included in IAL 1.1 because they use only 13 digits numbers.

On the other hand, there are other apps requiring more accurate authentication in Thailand. For example, in addition to the use of the 13-digit number, more secure approaches have been introduced in the Tang Rat app launched by the Digital Government Development Agency. It has four steps to identify a person as follows.

1) Take a picture of the national identity card to scan information, using optical character recognition technology.
2) The system check statements exactly, which is the same way to confirm the identity of the taxpayer.
3) Take a movie to compare the facial photo in the national identity card, using a high-reliability way of animation.
4) Set Password for requiring a login.

Hence, this certification level is included in IAL 1.3. From now on, it will be used as an innovation applying a digital platform to support government services via digital systems.

V. Comparison and Discussion

Comparing each context in both countries helped summarize the Pros and cons of the identification systems in Japan and Thailand. In Japan, the certification of residence is used to identify mainly and basically from a long time ago. It has two functions that register our addresses and certify themselves, going through changes along with historical problems. The government has distributed some profits per residence[6]. Now, it is in still time to adapt to changing the demands of society. That is why the individual number card is spread to prevent falsification more strictly.

The reasons the national individual card has not spread are not relating a function and an efficiency standpoint but political aspects. It is sometimes difficult to identify a person in terms of the used technologies. In the case of COVID-19 vaccinations, the required identification procedures are less rigorous because the government has to encourage the citizens to vaccinate.

On the other hand, in Thailand, the national individual card was issued to control the information of the population in 1909[7]. Currently, it is mainly used for identification in a variety of situations, associated with a national individual database.

TABLE IV
Pros and cons about identification system in Japan

Japanese online identification system	
Pros	Cons
· Easy to understand Japanese political contexts	· Need to apply for some identifications by ourselves
· Due to the World War II, we reluctant to have an individual card	· Possibilities of falsification
· a variety of identification cards	· Most of identification cards do not associate with national individual database

TABLE V
Pros and cons about identification system in Thailand

Thailand online identification system	
Pros	Cons
· The national individual card associates with national individual database	· The high risks of leaking individual informaiton
· Easy to obtain the identification card	· Most of services linked the specific card
· People tend to remember individual numbers	

As for the Pros, the identification system in Thailand improves the Japanese identification assignments. Moreover, as long as people remember their numbers, that might be helpful to distinguish them even in a disaster. In practice, the Mor Phrom app requires only individual numbers to authenticate. The government did not have to distribute new issues, unlike the Japanese vaccine ticket. Having a national identifying card for everyone promotes easily implementation of national policies.

VI. Reccomendations

As a result, concerning both countries' authentication, a national identity card should be spread like a Thailand one even in Japan. There are four required main factors in this card. The first factor is that the card should include an integrated circuit chip to prevent its falsification. The second is to associate it with a national database by using individual numbers. When governments implement policies, it works helpfully to record and control them. The third is that it should be easy to obtain and free to apply for, unlike a driver's license. Fourth is to attach a facial photo for certification with individual information on the card. According to the Japanese authentication, it found out to need to show a facial photo in the correct document to authenticate. Following these factors, the card will become the basis of certification and relate to a new online authentication system.

To show a real identity card by each time to identify leads to the possibility to leak individual information and the high risk of losing the card every time. For sake of avoiding that danger, the better authentication method is to introduce an identification digital ID shifted from physical certificates independent of each person, even though in person.

That is realized through an app that fits the requirements that it is possible to identify as long as entering OTP used phone number or e-mail and verifying by facial recognition or password after showing a real identification card only in the enrollment.

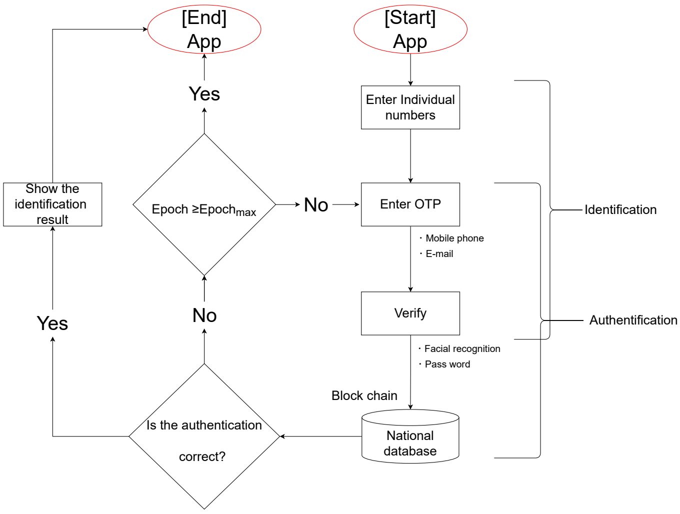


Fig. 3. The app identifying flowchart

The app connects national database storing individual information based on an identity card. The data are protected by blockchain technology, and also it records access logs, then sends a notification alert to users. Currently, there is a similar system in Thailand. It is called the National Digital ID platform(NDID). This system is available via the Bangkok Bank Mobile Banking application for Thai nationals. In the enrollment, people use citizen ID card information, OTP received via SMS and facial recognition. If they are going to certify, a digital process via NDID allows users to apply for particular services that request digital authentication, using facial recognition verification. A problem is the cost of introducing everywhere. They can not use this platform wherever they want to introduce it. It is important to think about how to save the cost and be available for every case.

VII. Conclusion and Further work

This paper seeks a better authentication method online, comparing the countries which have contrasting authentication systems in COVID-19. The ideal system should promote more accurate certification, avoid leaking individual information, be independent on a physical real card, be easy to introduce for everyone, and serve the same IAL levels in a variety of situations. There are two challenges to be addressed in implementation. One is that a national identity card must be spread widely first as realizing this system, however, Japanese encounter slow dissemination of a national identity card with an integrated circuit chip. Gaining an understanding of Japanese people is significant to solving this problem. The other is to develop a required app, saving cost to introduce it wherever it is needed indeed from now on. Through this app, it becomes possible to certify people online and offline via a national database even though they do not show a real identity card every time. That results in decreasing the risk of stealing individual information. The future authentication method could be reliable on an invisible digital thing.

References

[1] Tadashi Okada, Tadanobu Bandou, Motonobu Abekawa, Tashuhiro Yonekura, *The consideration about of public identification for proving*, IEICE Technical Committee, 2015, pp. 3-6.
[2] Hiroyuki Onishi, *The function and sysytem of identity cards in Japan*, Minpaku-tsushin 138, 2014, pp. 27.
[3] Igarashi Yasufumi, *The identity confirming system problem in Japan on Tohoku earthquake* , Fukuoka University, 2013, pp. 1-10.
[4] ETDA, *Reccomendation on ICT Standard for Electronic Transactions*, ETDA, 2021, pp. 2-10.
[5] Paul A. Grassi, James L. Fenton, Michael E. Garcia, *Degital Identity Guidelines*, NIST, 2017, SP 800-63-3.
[6] Tatsuo Chaya, *The historical consideration of the inhabitant resident's file and that change law*, Urban information systems institute, 2002, pp. 1-6.
[7] The international university, *Report on Research Study on the Current Status of National ID Systems in Other Countries*, Global communication center, 2012, pp. 61.

Advances in Information Systems and Technologies

AIST is a FedCSIS conference track aiming at integrating and creating synergy between disciplines of information technology, information systems, and social sciences. The track addresses the issues relevant to information technology and necessary for practical, everyday needs of business, other organizations and society at large. This track takes a socio-technical view on information systems and, at the same time, relates to ethical, social and political issues raised by information systems.

AIST provides a forum for academics and professionals to share the latest developments and advances in the knowledge and practice of these fields. It seeks new studies in many disciplines to foster a growing body of conceptual, theoretical, experimental, and applied research that could inform design, deployment and usage choices for information systems and technology within business and public organizations as well as households.

We call for papers covering a broad spectrum of topics which bring together sciences of information systems, information technologies, and social sciences, i.e., economics, management, business, finance, and education. The track bridges the diversity of approaches that contributors bring to the conference. The main topics covered are:

- Advances in information systems and technologies for business;
- Advances in information systems and technologies for governments;
- Advances in information systems and technologies for education;
- Advances in information systems and technologies for healthcare;
- Advances in information systems and technologies for smart cities; and
- Advances in information systems and technologies for sustainable development.

AIST invites papers covering the most recent innovations, current trends, professional experiences and new challenges in the several perspectives of information systems and technologies, i.e. design, implementation, stabilization, continuous improvement, and transformation. It seeks new works from researchers and practitioners in business intelligence, big data, data mining, machine learning, cloud computing, mobile applications, social networks, internet of thing, sustainable technologies and systems, blockchain, etc.

Extended versions of high-marked papers presented at technical sessions of AIST 2015-2021 have been published with Springer in volumes of Lecture Notes in Business Information Processing: LNBIP 243, LNBIP 277, LNBIP 311, LNBIP 346, LNBIP 380, LNBIP 413 and LNBIP 442.

Extended versions of selected papers presented during AIST 2022 will be published in Lecture Notes in Business Information Processing series(LNBIP, Springer).

- Data Science in Health, Ecology and Commerce (4th Workshop DSH'22)
- Information Systems Management (17th Conference ISM'22)
- Knowledge Acquisition and Management (28th Conference KAM'22)

Track Chairs

- **Ziemba, Ewa,** University of Economics in Katowice, Poland
- **Chmielarz, Witold,** University of Warsaw, Poland
- **Cano, Alberto,** Virginia Commonwealth University, Richmond, United States

Program Chairs

- **Chmielarz, Witold,** University of Warsaw, Poland
- **Miller, Gloria,** maxmetrics, Germany
- **Wątróbski, Jarosław,** University of Szczecin, Poland
- **Ziemba, Ewa,** University of Economics in Katowice, Poland

Program Committee

- **Ben-Assuli, Ofir,** Ono Academic College, Israel
- **Białas, Andrzej,** Instytut Technik Innowacyjnych EMAG, Poland
- **Byrski, Aleksander,** AGH University Science and Technology, Poland
- **Christozov, Dimitar,** American University in Bulgaria, Bulgaria
- **Dang, Tuan,** Posts and Telecommunications Institute of Technology, Vietnam
- **Dias, Gonçalo,** University of Aveiro, Portugal
- **Drezewski, Rafal,** AGH University of Science and Technology, Poland
- **Grabara, Dariusz,** University of Economics in Katowice, Poland
- **Halawi, Leila,** Embry-Riddle Aeronautical University, USA
- **Kania, Krzysztof,** University of Economics in Katowice, Poland
- **Kapczyński, Adrian,** Silesian University of Technology, Poland
- **Kluza, Krzysztof,** AGH University of Science and Technology, Poland
- **Kovatcheva, Eugenia,** University of Library Studies and Information Technologies, Bulgaria

- **Kozak, Jan,** University of Economics in Katowice, Poland
- **Ligeza, Antoni,** AGH University of Science and Technology, Poland
- **Ludwig, Andre,** Kühne Logistics University, Germany
- **Luna, Jose Maria,** University of Cordoba, Spain
- **Michalik, Krzysztof,** University of Economics in Katowice, Poland
- **Naldi, Maurizio,** LUMSA University, Italy
- **Nguyen, Thi Anh Thu,** The University of Danang, Vietnam
- **Pham, Van Tuan,** Danang University of Science and Technology, Vietnam
- **Rechavi, Amit,** Ruppin Academic Center, Israel
- **Rizun, Nina,** Gdansk University of Technology, Poland
- **Rollo, Federica,** University of Modena and Reggio Emilia, Italy
- **Rusho, Yonit,** Shenkar College of Engineering and Design, Israel
- **Sałabun, Wojciech,** West Pomeranian University of Technology, Poland
- **Santiago, Joanna,** Universidade de Lisboa – ISEG, Portugal
- **Sikorski, Marcin,** Gdank University of Technology, Poland
- **Solanki, Vijender Kumar,** CMR Institute of Technology(Autonomous), Hyderabad, TS, India
- **Taglino, Francesco,** IASI-CNR, Italy
- **Tomczyk, Łukasz,** Pedagogical University of Cracow, Poland
- **Webber, Julian,** Osaka University, Japan
- **Ziemba, Paweł,** University of Szczecin, Poland

28th Conference on Knowledge Acquisition and Management

KNOWLEDGE management is a large multidisciplinary field having its roots in Management and Artificial Intelligence. Activity of an extended organization should be supported by an organized and optimized flow of knowledge to effectively help all participants in their work.

We have the pleasure to invite you to contribute to and to participate in the conference “Knowledge Acquisition and Management". The predecessor of the KAM conference has been organized for the first time in 1992, as a venue for scientists and practitioners to address different aspects of usage of advanced information technologies in management, with focus on intelligent techniques and knowledge management. In 2003 the conference changed somewhat its focus and was organized for the first under its current name. Furthermore, the KAM conference became an international event, with participants from around the world. In 2012 we’ve joined to Federated Conference on Computer Science and Systems becoming one of the oldest event.

The aim of this event is to create possibility of presenting and discussing approaches, techniques and tools in the knowledge acquisition and other knowledge management areas with focus on contribution of artificial intelligence for improvement of human-machine intelligence and face the challenges of this century. We expect that the conference&workshop will enable exchange of information and experiences, and delve into current trends of methodological, technological and implementation aspects of knowledge management processes.

Topics

- Knowledge discovery from databases and data warehouses
- Methods and tools for knowledge acquisition
- New emerging technologies for management
- Organizing the knowledge centers and knowledge distribution
- Knowledge creation and validation
- Knowledge dynamics and machine learning
- Distance learning and knowledge sharing
- Knowledge representation models
- Management of enterprise knowledge versus personal knowledge
- Knowledge managers and workers
- Knowledge coaching and diffusion
- Knowledge engineering and software engineering
- Managerial knowledge evolution with focus on managing of best practice and cooperative activities
- Knowledge grid and social networks
- Knowledge management for design, innovation and eco-innovation process
- Business Intelligence environment for supporting knowledge management
- Knowledge management in virtual advisors and training
- Management of the innovation and eco-innovation process
- Human-machine interfaces and knowledge visualization

Technical Session Chairs

- **Hauke, Krzysztof,** Wroclaw University of Economics, Poland
- **Nycz, Malgorzata,** Wroclaw University of Economics, Poland
- **Owoc, Mieczyslaw,** Wroclaw University of Economics, Poland
- **Pondel, Maciej,** Wroclaw University of Economics, Poland

Program Committee

- **Andres, Frederic,** National Institute of Informatics, Tokyo, Japan
- **Berka, Petr,** Prague University of Economics and Business, Czech Republic
- **Bodyanskiy, Yevgeniy,** Kharkiv National University of Radio Electronics, NURE, Ukraine
- **Chomiak-Orsa, Iwona,** Wroclaw University of Economics and Business, Poland
- **Christozov, Dimitar,** The American University in Bulgaria
- **Chudán, David,** University of Economics, Prague, Czech Republic
- **Hernes, Marcin,** Wrocław University of Economics and Business, Poland
- **Jan, Vanthienen,** Katholieke Universiteit Leuven, Belgium
- **Kliegr, Tomáš,** Prague University of Economics and Business, Czech Republic
- **Kluza, Krzysztof,** AGH University of Science and Technology, Poland
- **Ligęza, Antoni,** AGH University of Science and Technology, Poland
- **Mercier-Laurent, Eunika,** Jean Moulin Lyon 3 University, France
- **Perechuda, Kazimierz,** Wroclaw University of Economics and Business, Poland

- **Schreurs, Jeanne,** Hasselt University, Belgium
- **Singh, Pradeep,** KIET Group of Institutions, Delhi-NCR, Ghaziabad, U.P., India
- **Singh, Yashwant,** Jaypee University of Information Technology Waknaghat, India
- **Sobińska, Małgorzata,** Wrocław University of Economics and Business, Poland
- **Stankosky, Michael,** The University of Scranton, USA
- **Tanwar, Sudeep,** Institute of Technology, Department of CE, Nirma University, Ahmedabad (Gujarat), India
- **Tyagi, Sudhanshu,** Thapar Institute of Engineering & Technology, India
- **Vasiliev, Julian,** University of Economics – Varna, Bulgaria
- **Zhu, Yungang,** College of Computer Science and Technology, Jilin University, China

Position Papers of the of the 17th Conference on Computer Science and Intelligence Systems pp. 99–106 DOI: 10.15439/2022F277
ISSN 2300-5963 ACSIS, Vol. 31

Open Data for simulation to determine the efficient management of parking spaces in Smart City

Iwona Chomiak-Orsa, Piotr Domagała, Andrzej Greńczuk,
Wojciech Grzelak, Artur Kotwica, Kazimierz
Perechuda Wroclaw Uniwersyty of Economics and Business, Komandorska 118/120, 53-345
Wrocław Poland;
E-mail:{iwona.chomiak-orsa, piort.domagała, andrzej.greńczuk,
wojciech.grzelak, artur.kotwica, kazimierz.perechuda}@ue.wroc.pl

Abstract—**The problem of optimization and effective management of parking spaces is one of the main problems faced by modern cities. Therefore, in creating Smart Cities solutions, more and more attention is focused on the possibilities of using advanced ICT tools to improve these processes. According to the authors of the article, such a method can be the creation of the so-called Digital Twin. In order to present the simulation possibilities that can be achieved by using a Digital Twin, the authors identified the chances of obtaining the data resources necessary for creating the simulation. Identification and evaluation of data sources had a character of a pilot study and referred to four car parks located in Wroclaw. On the example of one of them, an attempt was made to create an indicator of the type of traffic. The theoretical considerations and research presented in the paper are elements of research on creating Smart City solutions carried out by the team of authors.**

I. Introduction

NOWADAYS, we can observe the growing use of ICT (Information-Communication Technology) in various areas. ICT plays an important role in our lives and the proper functioning of organizations. It is used not only to facilitate communication, data storage, but more and more in every possible way. The current development of technology has focused on the use of mathematics, statistics and econometric models in analytics and decision-making processes. The development of tools and new technologies has resulted in the development of new tools that help reflect reality and make changes on it before they are implemented in "real life".

The created tools allow not only to model a potential scenario (through ongoing verification of "combined" elements), but also to recreate it in a virtual environment, change and assess the quality of these changes. Most people make decisions and/or create new solutions based on the simulation of a specific thing/phenomenon. The word "simulation" itself comes from Latin and means "pretending". It can also be said that simulations play a bigger role today than a few years ago. With the help of simple simulations, we can quickly and easily generate the output data, based on the input sample. Such results can then be further analyzed (including the use of data drilling algorithms) to detect appropriate relationships between the data and create specific models.

Simulations are used in various fields such as economics, including business, mathematics, computer simulation games, engineering sciences. Programs supporting the creation of simulations can be domain-specific as well as general use. This means that not every program will be able to implement a specific project.

The authors in this article continue their research in the field of efficient parking space management in Smart Cities. During the analysis of the literature, it was noticed that until now computer simulation was not used. A "simulation" was used, the aim of which was to create a prototype of the system. These prototypes either generated data based on a specific algorithm or tried to create a basis for a simulation model.

The article consists of the following parts. The first part discusses the role of simulation, Digital Twin and open data in improving Smart City processes. The second part deals with the problems of modern parking spaces in cities. The third part shows how open data can be used in the simulation of parking spaces in cities on the example of four parking lots in Wrocław. The whole thing ends with a discussion and conclusions in the area of open data usability in the simulation process.

II. Simulation based approach to streamlined Smart City processes

Simulation is an approximate reproduction of the phenomena or behavior of an object using its model. A special type of model is a mathematical model, often written in the form of a computer program. The concept of simulation was borrowed from traditional language learning. To "simulate" means to look similar to other people or to copy the behavior of such people. In the context of computer simulation, we are talking about copying the operation of the entire system or copying specific situations with the use of a computer program.

More precisely, "computer simulation" can be called a numerical method used to carry out experiments on specific types of mathematical models that characterize, using a digital machine, the operation of a complex system over an extended period of time.

From these definitions it can be concluded that:

- computer simulation is a method based on conducting research on dynamic models that discuss the existing or developed systems,
- the research reason for the computer simulation is to obtain information about the work of the analyzed system over time,
- computer simulation uses a computer program as a work tool while carrying out the research objective, which is the official representation of the model of the analyzed system.
- computer simulation as a method is a system of consciously selected research activities, ie the structure of phased works leads to the achievement of the set research goal.

Conducting a simulation enables the analysis of the process in various variants, which are verified in a virtual way, thus not affecting the activity of the process in real time. However, based on well-developed control parameters, consistent with the actual state, it can be said with high probability that the analyzed process variant has a chance to be implemented in the economic reality [7]. Each simulation requires the definition of basic principles [4]:

- in the case of complex processes subject to simulation, it is necessary to properly select the tool used for the simulation and detailed modeling of the parameters of the analyzed process and the system in which it operates, defining the input data and defining the goal,
- in the case of flexible processes subject to simulation, it is necessary to frequently change the values of the control parameters,
- basing the analysis on average values of parameters carries the risk of misinterpretation,
- the simulation must be done in a timely manner to obtain the greatest benefit.

The simulation model design procedure includes the following stages [9]:

- identification of the simulated object using one of the two approaches: top-down, in which the main process is detailed into sub-processes and activities; bottom-up, which starts with defining all activities, and then grouping them into sub-processes and main processes,
- developing diagrams of the simulated process using IT tools (the number of hierarchy levels depends on the detail of the analyzed process),
- collecting input data and parameters, and then entering them into the simulation model,
- model verification, which comes down to comparing the behavior of the simulation model with the actual behavior of a given system (Figure 1).

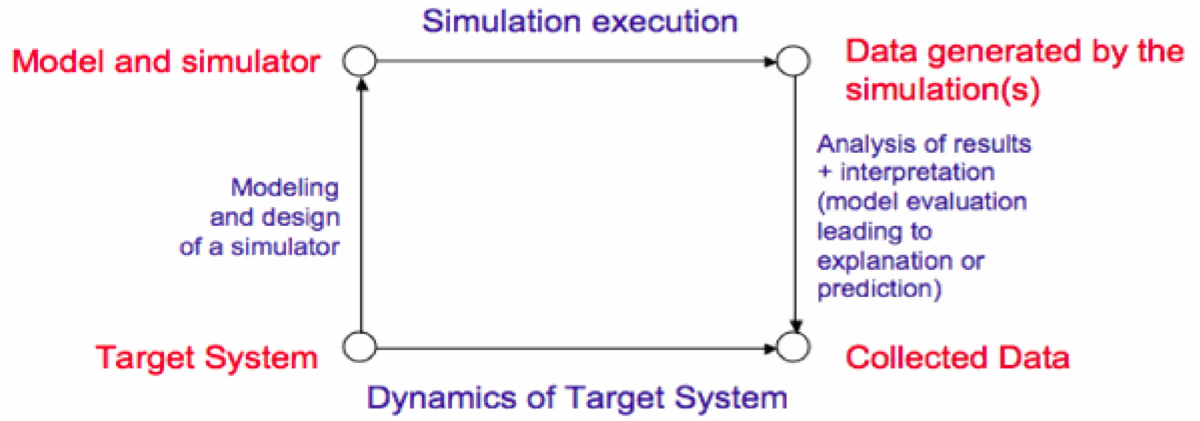

Simulation execution
Model and simulator
Data generated by the simulation(s)
Modeling and design of a simulator
Analysis of results + interpretation (model evaluation leading to explanation or prediction)
Target System
Collected Data
Dynamics of Target System

Figure 1. A general schema describing the usage of simulation as a predictive or explanatory instrument. Source: (Bandini, Manzoni, Vizzari,

Computer simulation also allows you to extend the operating time of the system, because it can be used to examine the detailed structure of changes that could not be observed in real time.

A Digital Twin is a mirror image of a physical process that is articulated alongside the process in question, usually matching exactly the operation of the physical process which takes place in real time (Batty, 2018). The term Digital Twin denotes a replica of a physical asset, process or system used for control and decision making [10].

Digital Twins are adopted by several disciplines. They have been applied to agriculture [8], Industry 4.0 in the context of smart manufacturing [5], prediction of the ergonomic performance in automotive industry Caputo, 2019.

A Digital Twin is expected to enhance city management and operations to achieve a smarter and sustainable city and a higher quality of life for its citizens. First implementations of Digital Twins in context of Smart Cities have arisen. In Zurich the city Digital Twin enhances city administration and support urban planning decision-making processes (Figure 2). To enable the use of the Digital Twin, open governmental data is being utilized in order to facilitate contributions from the different stakeholders and their accessibility to the city data [11].

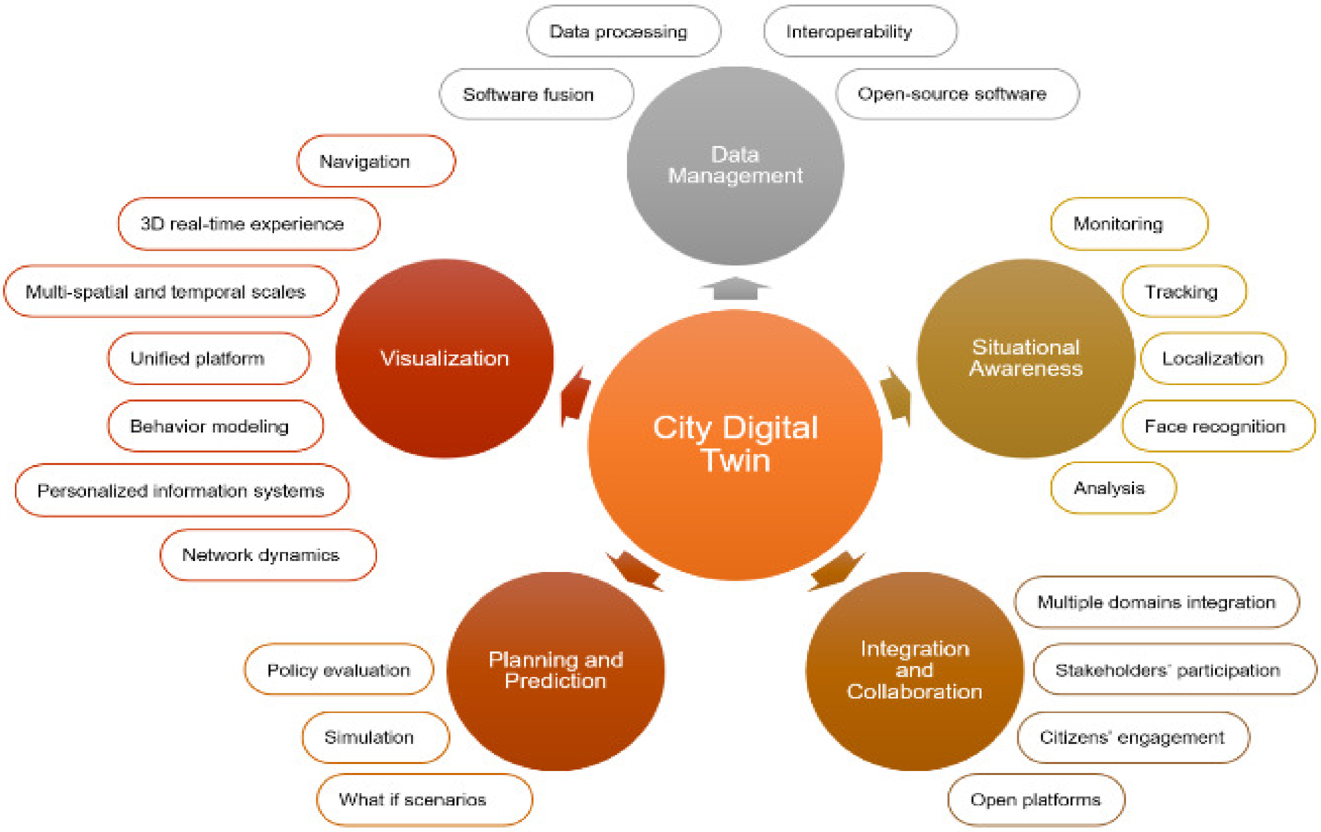


Figure 2. City Digital Twin potential. Source: [11]

Digital Twin is increasingly being explored as a means of improving the performance of physical entities through leveraging computational techniques, themselves enabled through the virtual counterpart [6]. Digital Twin very often starts life as a Digital Twin Prototype and helps in modelling, testing, optimization of a real product or asset. In its essence Digital Twin enables the application of a knowledgeable, data driven approach to the monitoring, management, and improvement of a product or a city asset throughout it's life-cycle. A digital twin can be perceived as an opportunity to enhance city planning and operability. Digital Twins enable performing simulations on the virtual model. Forecasting and optimization of the physical entity's performance are realizable, and thus the optimization of the physical counterpart's performance can be achieved. Digital Twins can also engage the citizens in creating new plans for the city and enhancing public decision-making.

Data describing real city processes is a key resource required to build and maintenance a relevant implementation of a Digital Twin that delivers reliable insights. Proper data availability is the one of most significant challenge in Smart City Digital Twin area. We can distinguish the following issues regarding data:

- lack of open data sources in specific cities referring to a selected city domains,
- large-sized, complex, and heterogeneous nature of the city data,
- lack of a widely accepted standards for the data models and design schemas to facilitate the development of the city models.

Open data is data that anyone can access, use and distribute. This definition, formulated by the Open Data Institute [12], can be applied both to public data (generated in the public sector, e.g. by government administration or other state institutions) and to research data. As with scientific publications, also with data there are technical and legal barriers that must be removed in order for data to be considered open. Two stack models to assess the degree of data openness can be distinguished:

- The FAIR model (Findable, Accessible, Interoperable, Reusable) is a set of recommendations formulated by a group of experts from the FORCE11 organization, which should guide people opening research data [13]. This model identifies four most important aspects of open data: it should be well searchable, accessible, interoperable and reusable. The authors strongly emphasize that due to the way data is used in the modern world and in modern science in particular - data should be available both in a human-readable form and in a form suitable for machine analysis.
    - The 5-star open data model was developed primarily with public open data in mind, but can be applied to research data [14]. In this model, the removal of legal barriers means that the data is awarded one star, the next four stars relate to the removal of technical barriers to the use of data.

Open (government) data is information collected, provided, or paid for by public authorities (also known as public sector information) that is made freely available and re-used for any purpose.

Problems using open data:

- Purpose - open data allows you to look deeper into a specific topic that we want to know more about. Economic operators can also use open data to refine their customers' profiles and better adapt to their needs. Whether used for private or commercial use, open data offers many possibilities.
- License - the license allows the use of data in a way that interests us (eg if we create a commercial application, is it allowed to re-use the data in a commercial way). The license may require you to identify the data publisher when used, i.e. we must provide the data owner when we make the product or service available. This requirement is called attribution.
- Data format - when we find that a specific data set contains exactly what we need, we can download it in one of the available formats. Based on our IT knowledge, we select the most appropriate type of file. The most common format for tabular data is ".csv". It allows you to add information to a file and perform calculations using the data contained in it. Data sets that can be changed are published in an open format. Most datasets are available in an open format, but please note that some formats (e.g. ".pdf") are not changeable.
- Data quality - the page from which we want to download the data set should contain the date of the last modification of the file. If we need data from a specific period, it is necessary to check whether the information about the time period is provided or whether the file has been recently updated. You should also make sure that the information you expect to find in the file is actually included in it and that you recognize individual labels.

Checklist developed by the Open Data Institute [15]:

- form:
    - how is the data processed?
    - are they in a processed or unprocessed form?
    - how will the form affect the analysis / product / application?
    - what syntactic (language) and semantic (meaning) transformations will be required?
    - are they compatible with other, already owned data sets?
- quality:
    - how up-to-date is the data?
    - how regularly are they updated?
    - are all fields and data context understandable?
    - how long will they be published? what is the publisher's commitment?
    - what do we know about data accuracy?
    - how is the missing data problem solved?

Open public data is the data of institutions and offices that anyone can use. On the basis of open public data, more and more modern products and services are created in Europe and around the world. Open data is a real source of real savings in money and time for administrations and citizens. Citizens, including entrepreneurs, can use public data resources to pursue their own goals, developing their business or research.

One of the conditions for digital development is quick and effective access to high-quality data, which allows for the creation of more innovative solutions, e.g. in the area of the so-called artificial intelligence, or to put it more precisely, automation and prediction. The European Data Strategy [14], which was developed by the European Commission, identifies the openness of high-quality data and the value of data as one of the pillars of building the competitiveness of the European Union economy. Therefore, in 2021, the adoption of implementing acts was planned, which will enable the public sector to make data sets available in a machine-readable format and via application programming interfaces (APIs). Data openness is indicated as a key element to stimulate innovation in many sectors of the economy, but also in science, and machine learning technologies, natural language processing or the Internet of Things require increasing the supply of the said data [16].

III. NEW PARKING REQUIREMENTS

Environmental friendly transport system is one of the most significant parameters of the smart city. There is some kind of the tacit contradiction between dynamic (car movement) and static (car parking) aspects of the traffic in contemporary European cities, which are more and more closed for private and business transport. In the European Union won the static concept, which means that the big transport vehicles cannot enter city closed system; they must park on the city boarder. Also private car movement is strongly oriented toward city parking system. The real transport is more oriented toward city outside environment; in the city it is totally restricted to minimum with especially for gasoline vehicles.

The EU ecology strategy and environmental regulations are mainly oriented toward:

- significant reduction of CO2 emission,
- development of environmental friendly energy sources,
- enhancement of electric and hybrid car production.

One of the real determinant for such EU development is obviously actual geopolitical situation connected with diversification of the global energy sources. These, above mentioned principles, create more sophisticated expectations for:

- car industry,
- transport system within EU and
- smart city models.

On the base of EU ecology model we can identify the following new smart city problems:

- reduction to minimum the quantity of gasoline vehicles within cities,
- supporting and extension of the city routes fort electric bikes, scooters, motorbikes, autos and small trucks,
- renovation of existing and construction of the new parking with electric loading systems.

We notice in the last years significant grow of the global electric cars sales volume especially of the following car companies:

- Toyota,
- KIA,
- Tesla and
- European producers.

This trend is inevitable; e.g. Volkswagen and other German car manufacturers invest strongly and construct new auto electric battery factories. The lack of electric battery loading stations creates another, big problem for electric car users. Therefore the authors of this paper suggest analyses and optimize of city parkings in the context of:

- reconstruction of the existing parkings toward some kind of "electric parking plant", where the car drivers can load their electric vehicles,
- construction of the new parkings, fully equipped in the modern electric loading systems.

Such investments will give for parking owners significant competitive advantage.

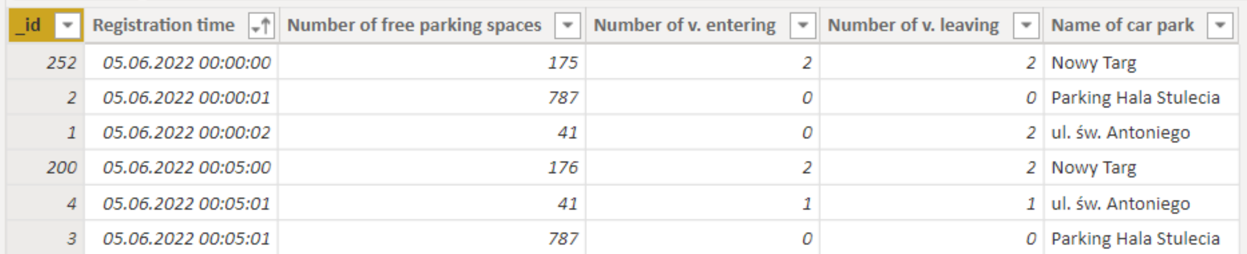
IV. OPEN DATA IN SIMULATION OF PARKING SPACES ON THE EXAMPLE OF WROCŁAW

RESARCH METHOLOGY

The aim of this research is to find simply measure of character of parking traffic. To achieve the aim was using data gathered from Wroclaw city portal.

The data for the study come from the website of the city of Wrocław, https://www.wroclaw.pl/open-data/dataset/zapelnienieparkingowodczytza48h_data. The data is open data and collected from 4 car parks in Wrocław, Hala Stulecia, National Forum of Music, Nowy Targ and St. Anthony. St. Anthony is the subject of our research. The data is published on the website every 48 hours, which means adding more data.

Table 1. Structure of data from the website of the city of Wroclaw



_id	Registration time	Number of free parking spaces	Number of v. entering	Number of v. leaving	Name of car park
252	05.06.2022 00:00:00	175	2	2	Nowy Targ
2	05.06.2022 00:00:01	787	0	0	Parking Hala Stulecia
1	05.06.2022 00:00:02	41	0	2	ul. św. Antoniego
200	05.06.2022 00:05:00	176	2	2	Nowy Targ
4	05.06.2022 00:05:01	41	1	1	ul. św. Antoniego
3	05.06.2022 00:05:01	787	0	0	Parking Hala Stulecia

Shared public data is composed of six columns. The structure od the file is presented in **Błąd! Nie można odnaleźć źródła odwołania.**3.

The indicator of a type of parking traffic was defined as a deference between number of cars income into parking and number of cars goes outside of the parking in the same period of time

Type traffic=number of enters cars - number of leaving cars

That defined indicator allows for the study of the character of parking traffic. When Type traffic is positive that means, that entering traffic is higher than leaving traffic and the number of parking spaces reduce. In case when Type traffic is negative that means, that entering traffic is lower than leaving traffic and the number of parking space increase. The sign of Type traffic determines the direction of parking traffic in the time period, while the value of Type traffic talks about the intensity of the direction of traffic.

That simply measure could be useful especially with combination with the data visualization technique.

To visualize the Type traffic, first of all, the original parking data must be transformed into new structure. The model of data show **Błąd! Nie można odnaleźć źródła odwołania.**2. To transform the model of parking data was use Python programming language and the Pandas package.

Table 2. Transformed structure of the dataset.

Column's name	**Description**
Id	Record id
Id parking	A number representing the car park name
Register time	Time of register the event. The time format is hour:minutes
Day number	A number of the day of the week
Number of entering cars	A number of vehicles entering the parking. Value is a sum of enter cars in period of 15 minutes
Number of leaving cars	A number of vehicles leaving the parking. Value is a sum of leaves cars in a period of 15 minutes

It seems that the visualization of the Type traffic indicator in 15th minutes periods is too high, and this is the reason why records were aggregated to one hour period. The new model of data is presented in Table 3.

Table 3. New format parking data with Type traffic indicator

Column's name	**Description**
Id	Record id
Id parking	A number representing the car park name
Date	Date of register the event. The format is year-mont-day
Day number	A number of the day of the week
Type traffic	Difference between the number of input cars and the number of output cars in one hour period

1. Files with data were downloaded every 2 days and stored in shared drive to collect data covering 2 days of car park functioning.
2. Rows were merged to generate 1 consistent data set presenting car park functioning through 2 days.
3. Data was imported to Microsoft Power BI Desktop.
4. Preparation of data visualisations in Microsoft Power BI Desktop to identify possible car park deficiencies.
5. Transformation of original file int a structure enabling a comparative analysis between events of every day. The structure is presented in Table 4.
6. Data analysis in Python language programming with Pandas package.
7. Preparation of data visualisations in Python language with Plotly package to identify possible car park deficiencie

Table 4. Transformed structure of the dataset.

Column's name	**Description**
Id	Record id
Id parking	A number representing the car park name
Register time	Time of register the event. The time format is hour:minutes

Day number	A number of the day of the week
Number of entering cars	A number of vehicles entering the parking. Value is a sum of enter cars in period of 15 minutes
Number of leaving cars	A number of vehicles leaving the parking. Value is a sum of leaves cars in a period of 15 minutes

RESULT

Figure 3 shows the number of vehicles entering the parking lot from 4th of April to 7th of April, and Figure 5 shows the number of vehicles leaving the parking lot.

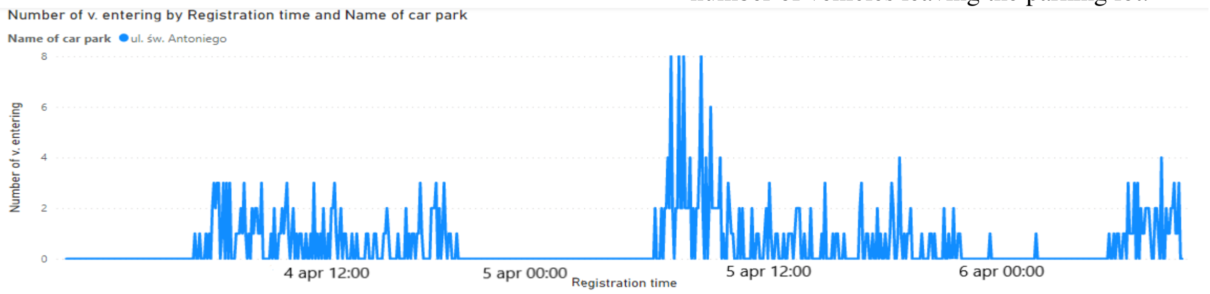


Figure 3. Number of vehicles entering the St. Anthony

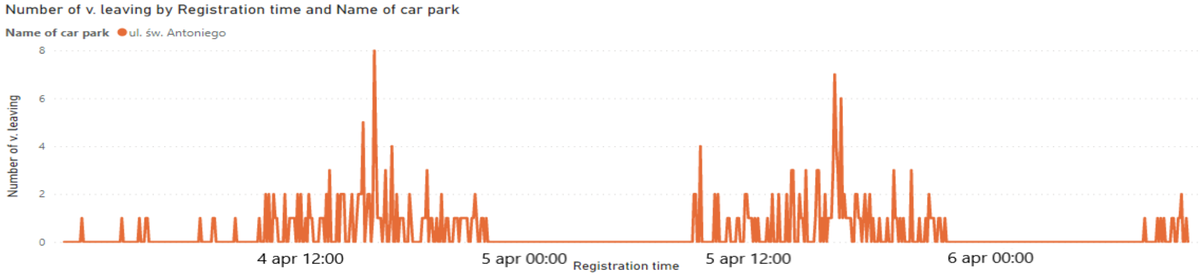


Figure 4. Number of vehicles leaving the St. Anthony

A sum of the number of entering cars and the sum of the number of leaving cars in an hour period should let for better understanding visualization of the Type traffic indicator.
To visualise date Plotly package was used. In the
Figure is presented visualization of the Type traffic in one hour period. Thera are data comes from four days from 4th of April to 7th April

Figure 5 is presented a visualization of the Type traffic in a 24hour period from 0 a.m to 23 p.m. The X-axis represents one-hour time periods. Every period shows an aggregated value of Type traffic. There are data comes from four days from the 4th of April to the 7th of April.

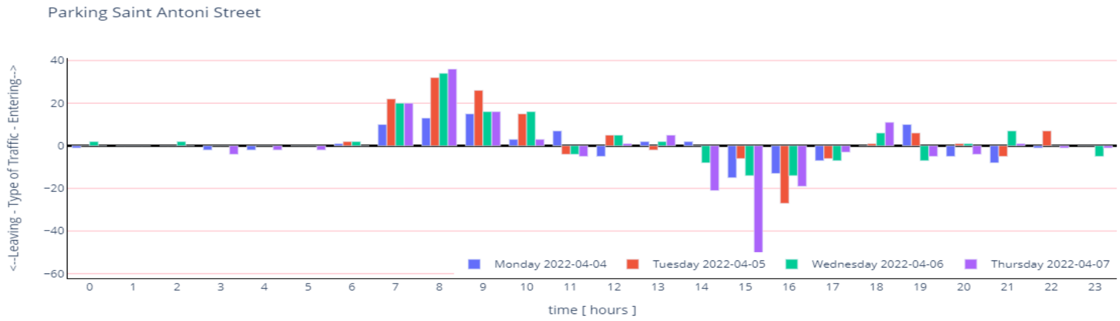


Figure 5. Visualization of the Type of traffic

Analysing the graph below it is clear, that Type traffic in period from 6 am to 11 am has a positive value, which means entering traffic is higher than leaving traffic, and free parking spaces reduce. The character of parking traffic is changing from 2 pm to 6 pm, where Type traffic is negative, which means parking spaces increase. The same character of parking traffic repeating in every day.
The Type traffic indicator could be a useful measure to describe the character of parking traffic.

V. Discussion

The aim of this paper was to highlight the possibilities offered by intelligent IT tools in the area of simulating selected processes occurring in urban space.

Authors decided to indicate possibilities of using Digital Twin method in analysis, simulation and support of intelligent solutions related to the management of parking space. Due to the pilot character of the presented research, data obtained from open sources made available by Wroclaw were used in the analysis.

At this stage of designed research, the authors diagnosed a significant problem in gaining access to data from different cities in Poland. In the initially defined research procedure, it was planned to contrast data on functioning of parking lots from selected, provincial cities in Poland. Unfortunately, except for Wroclaw, the authors were not able to obtain credible, reliable and up-to-date data on occupancy of urban parking lots in other voivodship cities. It should be pointed out that especially the parameter of actuality and systematic data refreshing is important for the credibility of created simulations. That is why, the authors decided to present a study using data made available by Wroclaw, which are updated every 10 minutes and concern 4 city parking lots.

The conducted research has shown that the availability of open parking lots allows you to conduct an analysis in terms of the occupancy of parking spaces, as well as to observe various types of anomalies related to it. What they lack is a link to other data to identify this fact. However, it is possible through additional data acquisition, e.g., about events taking place in the vicinity of a given car park or in the city. These analyzes will allow you to build a specific model, as well as determine the conditions for its functioning. As an example, we can point to, for example, the simulation of store queues, which can be used to determine the potential behavior of customers in the plotted conditions. In the field of parking space management, it will be possible to determine the occupancy of parking spaces, as well as, to a further extent:

- creating additional applications informing which parking space is free,
- which seats will be occupied at a certain point in time,
- total occupancy of the car park.

In addition, if additional metadata such as location, city, nearby public facilities are used, it may be possible to accurately plan and build additional/new parking spaces.

VI. Conclusion

Both visual analyses show us, that the transformation original dataset allows us to better understand parking traffic. The moments of intensive traffic are presented. In such moments traffic related issues occur, that Digital Twin and simulation can address. We can observe e.g the repeating distribution of the number of entering cars and also unexpected traffic caused by external factors like public events. Shared datasets by the Wroclaw are good datasets to start the process of parking traffic analysis. Having such files it is possible to start building a simulation addressing the issue.

References

[1] Bandini, S., Manzoni, S., Vizzari, G. (2009). Agent Based Modeling and Simulation: An Informatics Perspective, Journal of Artificial Societies and Social Simulation 12 (4).

[2] Batty, M. (2018). Digital twins. Environment and Planning B: Urban Analytics and City Science, 45(5), 817-820.

[3] Caputo, F., Greco, A., Fera, M., & Macchiaroli, R. (2019). Digital twins to enhance the integration of ergonomics in the workplace design. International Journal of Industrial Ergonomics, 71, 20-31.

[4] Dullinger, K.H. (2009). Simulation in der Logistik - neue Anwendungsfelder. LogForum, Vol. 5, Issue 3, 1-12.

[5] Jiang, Y., Yin, S., Li, K., Luo, H., & Kaynak, O. (2021). Industrial applications of digital twins. Philosophical Transactions of the Royal Society A, 379(2207), 20200360.

[6] Jones, D., Snider, C., Nassehi, A., Yon, J., & Hicks, B. (2020). Characterising the Digital Twin: A systematic literature review. CIRP Journal of Manufacturing Science and Technology, 29, 36-52.

[7] Koliński, A., Śliwczyński, B., Golińska-Dawson, P. (2019). Wykorzystanie symulacji jako narzędzia wspomagającego proces oceny efektywności produkcji w przedsiębiorstwach produkcyjnych. E-mentor nr 3 (75) / 2018.

[8] Pylianidis, C., Osinga, S., & Athanasiadis, I. N. (2021). Introducing digital twins to agriculture. Computers and Electronics in Agriculture, 184, 105942.

[9] Rodawski, B. (2006). Simulation of logistics processes. LogForum, Vol. 2, Issue 1, 1-15.

[10] Vatn, J. (2018, June). Industry 4.0 and real-time synchronization of operation and maintenance. In in Proceedings of the 28th International European Safety and Reliability Conference (pp. 681-5).

[11] Shahat, E., Hyun, C. T., & Yeom, C. (2021). City digital twin potentials: A review and research agenda. Sustainability, 13(6), 3386.

[12] www1 https://theodi.org/what-is-open-data (date of access: 15.04.2022).

[13] www2 https://www.force11.org/group/fairgroup/fairprinciples (date of access: 17.04.2022).

[14] www3 https://www.parp.gov.pl/component/content/article/71119:otwarte-dane-filarem-innowacyjnej-gospodarki-co-przyniesie-nowa-ustawa#_edn1 (date of access: 15.04.2022).

[15] www4 http://theodi.org/guides/the-open-data-consumers-checklist (date of access: 17.04.2022).

[16] www5 https://data.europa.eu/pl/trening/what-open-data (date of access: 15.04.2022).

DOI: 10.15439/2022F284
ISSN 2300-5963 ACSIS, Vol. 31

Digitalization impact on higher education – potential and risks

Beata Butryn, Katarzyna Hołowińska, Małgorzata Sobińska
Wroclaw University of Economics and Business
ul. Komandorska 118/120, 53-345 Wroclaw, Poland
Email: {beata.butryn, katarzyna.holowinska, malgorzata.sobinska}@ue.wroc.pl

Laura Martini
University of Padua,
Via VIII Febbraio, 2,
5122 Padova PD, Italy
Email: laura.martini0512@gmail.com.

Abstract—**The topis of education is often superseded by business, mainly because business is associated with rapid technological advancement and often brings lucrative profits. Perhaps this is why significant paradigm changes in the education field take place only during crises or exceptional situations that force a change of the current approach. So it was this time, that the pandemic of covid- 19 revealed how much education systems are not enough adapted to the technologically changing world. The main goal of the article is to present a general overview of the process of digitalization in the higher education field. The introduction part highlights the importance of today's education. The second part of the article describes the characteristic of digitalization in the context of education. The third section indicates the potential and risks of digitalization. The next part of the article points out the use of digitalization in the didactic process.**

I. Introduction

UNIVERSITIES, like businesses, are subject to constant pressure from the environment and are forced to take up challenges dictated by the changes taking place in the environment and actions that will ensure that their existence will not be endangered. Currently, the key challenges include such support for knowledge management/knowledge management within universities, thanks to which universities will be perceived as attractive for both candidates for studies (in the era of globalization of candidates increasingly often coming from all over the world) and organizations that will employ graduates such universities. On the other hand, universities should provide the best possible working conditions for their staff, including research, teaching, and administration staff. In this article, the attention will be focused primarily on the changes that have occurred and are taking place in relation to didactic processes, which consist in the virtualization and digitization of education processes.

Teaching processes are at the same time one of the key areas of university functioning and a critical element of the university's knowledge management system based on knowledge sharing.

The aim of the article will be an attempt to assess the current situation of higher education in the field of digitization and virtualization of processes, and above all a discussion of the purposefulness, potential and benefits of such changes, as well as the risks and pitfalls that may not be perceived in the light of contemporary trends and sometimes too thoughtless focusing on technological aspects / IT tools. This article will cover a critical analysis of the phenomenon of digitization of higher education institutions as a process that is undoubtedly necessary to meet the expectations of stakeholders (employees, students, business, etc.) and at the same time difficult to define measurable benefits.

The following research methods were used in the article: literature review, observations, and own experiences as well as interviews with randomly selected employees and students at Wroclaw University of Economics and Business. The interviews were conducted in the period February-May 2022, in the period after the pandemic, when the classes were already conducted in the stationary version.

II. Characteristic of digitalization

Digitalization is the process of converting information such as texts, pictures, or sounds into a digital format, that can be processed by a computer. In addition, digitalization means improving an organization's core business operations to satisfy customer requirements efficiently by the use of data and technology. In the education industry, the target customers can be students, teachers, staff, and alumni, and digitalizing the education sector can bring benefits to both students and faculty.

Digital transformation in education comprehends, among the many, the following tools and resources:

- AI Chatbots: A chatbot is a computer program that simulates human conversation through voice commands or text chats or both. Chatbot, short for chatterbot, is an artificial intelligence (AI) feature that can be embedded and used through any major messaging application.
- Adaptive Learning: is a technique for providing personalized learning, which aims to provide

efficient, effective, and customized learning paths to engage each student.

- Smart Classroom is an EdTech-upgraded classroom that enhances the teaching and learning process for both the teachers and the students by inculcating audio, video, animations, images, multimedia, etc. This increases the engagement factor and leads to better-performing students.
- Remote Proctoring: allows students to take an assessment at a remote location while ensuring the integrity of the exam. These systems require students to confirm their identity, and, during the exam, the system monitors students through video, looking for behavior that could indicate cheating.
- Video conferencing for online studies: Telecommunication in the form of a videoconference.
- AR/VR for a better learning experience: Augmented reality (AR) adds digital elements to a live view often by using the camera on a smartphone.

Digitization in recent years has been increasingly evolving and having an increasing impact on society, work, school, and people's lives. Furthermore, the Covid-19 pandemic has undoubtedly accelerated some of the digitization processes, with the aim of moving the economy, education, jobs, and relationships between people forward. In some nations, like Italy, there has been a lockdown, during which, without digital means, many activities both work and school would have been interrupted. Here smart working and distance learning has certainly played a key role. This has involved a reorganization of activities, such as the purchase of the computer and additional training on the use of the devices. With regard to the development of digitization in education in particular, several tools and resources, in general, have been employed, to enhance the student experience in particular. Among them we find some communication technology platforms, such as Zoom Video Communications, Google Meet, Microsoft Teams, and Skype, which allow, through video calls, to hold online classes, create virtual rooms, use support chats, and many other features. On these platforms and others, such as Classroom, there are additional possibilities of using, for instance, sharing of teaching materials, the electronic register, which also allows a direct link between teachers and students' families, and the possibility for students to register for admission or to exams via the app, the recording of progress of students during their course of studies. In addition, there are many other tools that provide a wide array of online learning options, such as web pages or YouTube channels with interesting points and explanatory videos or examples. Other useful tools, mainly in university settings, are digital libraries, i.e., archives, in which students can share their projects and work that can be viewed by all those enrolled in the same school. Of course, the use of this tool must be controlled so as to avoid any form of plagiarism. With this, students are motivated and challenged to be active and interact with each other constructively.

All of this has occurred starting with students in school and university, with online classes, some live, some deferred, with teachers and professors implementing new teaching techniques, sometimes more interactive. There is the possibility to create quizzes or online games for reviewing a given topic, for the benefit of students, and to help teachers with possible modifications of teaching intervention. Among these is Kahoot, an application that allows the creation of ad hoc quizzes with questions and answers of various kinds in a simple and effective way to make the lesson interactive and engaging.

Advantages

Digitalization brought some positive aspects, such as increased productivity and efficiency, especially from the economic point of view: fewer expenses, and less "wasted" time. In addition, people became more proficient in the use of electronic tools, with advantages also for their future, had the ability to find material online with considerable ease and to communicate despite distance and other problems, including indisposition at the health level. Another advantage of digital tools is their environmentally sustainable aspect. For example, the use of e-books, tablets, or laptops saves large amounts of paper and the ability to have it with you at all times.

Disadvantages

In addition to the many positive aspects of digitization, there are also some disadvantages. Among these, we find the difficulty of teachers and professors in keeping students' attention and respect alive, who certainly have more sources of distraction at home. Students have missed almost two years of "regular" school, in the classroom, with their peers and friends, so in many cases, their health and psyche have been negatively affected. This type of teaching also created financial difficulties in some families who did not have the ability to purchase a laptop. Furthermore, these means and resources, sometimes, can be used in the wrong way, for instance when students, find material easily, just "copy and paste" what they found online, without filtering the seriousness of the sources and without putting their creativity into it, thus leading them to a loss of reading and writing skills and critical sense. The worst thing they can do is verified when they take advantage of the situation, certainly in the wrong way, and, for example, in the case of an online test, have an outside person take it.

So, as mentioned earlier, the pandemic brought the above-listed positive and negative aspects. Since there was a certain urgency in finding adequate and suitable solutions for the period, some disadvantages and damages were inevitable. With prior experience, over time, by implementing new strategies, the problems related to digitization that occurred during the pandemic can be avoided or reduced, enhancing the positive aspects and advantages it brings.

III. POTENTIAL AND RISK OF DIGITIZATION OF DIDACTIC PROCESSES

The topic of teaching virtualization is not new, although during and after the COVID-19 pandemic, activities related to the implementation of diverse types of strategies and projects to virtualize university operations have certainly gained in importance and speed. Some aspects related to virtualization/digitization in education are highlighted by, among others: P. Petrov, M. Radev, G. Dimitrov [Petrov et al., 2022], M. Cuypers [Cuypers 2012], T. Pfeffer [Pfeffer, 2003] or K. Prosyukova, F. Shigapova [Prosyukova & Shigapova, 2020].

The issues gained importance with the emergence of the pandemic and the need for a rapid transition to distance learning. It also turned out that such a need for a moment gave an impulse for a very quick adaptation of employees and students to the new, remote form of work/teaching and learning and the university authorities were considering the next steps towards virtualization and digitization not only of teaching but also other areas of university functioning. There is an ongoing discussion involving the authorities of many universities in Poland on the digitization of Polish universities, both in the context of developing the educational offer and scientific and research cooperation and in relation to the marketing and recruitment strategies of Polish universities [Tytuła, 2020]. Also, foreign universities have been observing this phenomenon for years, examining the possibilities and introducing changes aimed at the best adjustment of the educational offer to the market expectations. At the same time, the limitations and weaknesses of changes consisting in digitization and virtualization of "teaching services" are also noticed. For example M. Cuypers, who analysed a potential relation between the emphasis on virtual web-based processes in an institution of higher education and the paradigm of internationalization comparing three German universities, a campus-based university, a mostly paper-based distance-learning university and a fully virtual distance-learning university in terms of their services and procedures, states that "campus-based university has advantages for the sociocultural and political rationales of internationalization due to the emphasis on face-to-face communication and on-campus services, while the virtual university succeeds for the educational and economical rationales of internationalization because of the more wide-spread influence of web services and timeless availability of content" [Cuypers, 2012].

As described in the earlier publication of the authors [Binsztok in. 2022] the last two decades have been a time when information and communication technologies have an increasingly greater and greater impact on all sectors of the economy, including the higher education sector, which, like business, is beginning to undergo gradual digital transformation. The environment expects universities, on the one hand, to conduct teaching processes in a way that will prepare students for the challenges of the modern world, and on the other - to create knowledge corresponding to contemporary social requirements and phenomena. Current students and applicants for studies economic studies are mostly people with developed digital skills, expecting modern forms and channels of interaction with lecturers or, more broadly, university employees.

Increasingly, applicants for studies are people from different countries, who expect certain standards in terms of education programs and tools used at universities. This is certainly the reason it is worth taking care of the quality of education and looking for ways of acquiring, codifying, and disseminating knowledge as well as conducting research at the highest scientific level. Well-selected and implemented solutions based on the latest ICT technologies and processes of digitization of knowledge processes appear here as a potential help in the pursuit of maximizing the level of organizational (domain) knowledge of universities.

However, implementing technological tools at universities is a major challenge and requires a systemic approach. Universities that consciously approach digital transformation must consider many factors, including increasing the digital competencies of all university employees, including academics (teachers, and scientists), using tools to support and develop didactic innovations, as well as conducting scientific activities or building relationships with students and graduates through new communication channels.

The following barriers/obstacles to the transition to modern forms of education on the part of universities can be noticed (own study based on [Binsztok i in., 2022]):

- low digital competencies of teachers resulting from poorly conducted campaigns promoting modern teaching tools,
- lack of motivation to use such tools,
- lack of or too limited training,
- lack of time to participate in trainings due to excess duties (didactic work plus scientific work; often the need to work over a working age),
- insufficient infrastructure,
- fear of excessive "bureaucracy" in the teaching process,
- academic teachers' fear and reluctance to limit the freedom in the selection of means and methods of working with students,
- employees' fear of excessive control by the employer,
- employees' fear of interference with privacy,
- the need to employ auxiliary / technical staff who would support educators in using specific tools, and thus - additional costs of introducing such a change.

On the part of students, such barriers may be:

- lack of motivation,
- lack of involvement in classes,
- students often hidden behind cameras do not actually participate in classes, although in the application they are shown as "active",

- limiting the time for discussions / brainstorming,
- limiting situations requiring critical thinking, looking for arguments for and against,
- limiting creativity by performing prepared / standardized tasks / simulations, etc.,
- overabundance of tools with which they come into contact during a brief period of study, e.g., in fields more closely related to IT,
- insufficient infrastructure (no laboratories) that would allow improving the ability to use specific tools (e.g., simulations) outside of class hours.

Students already, as promoters of bachelor's and master's theses have noticed, have problems with both verbal and written communication. It is exceedingly difficult to prepare a written expression that is longer than 1-2 pages. It seems that the short messages and commands used in communication with devices have impoverished the language of young people, so from this point of view, limiting "normal" interpersonal contacts in favour of virtual ones seems to be risky. At this point, it is also worth paying attention to health problems resulting from the lifestyle to which the ubiquitous technology and related "customs" accustom us. More people notice distinct types of back pain, problems with eyesight, hearing (people using headphones) or even mental problems resulting from the lack or insufficient number of contacts with another person or, for example, not keeping up with the news / not coping with a brief time of many new skills and knowledge.

Nevertheless, the digital transformation of higher education seems to be an irreversible process expected by all stakeholders of institutions making up this sector. The authorities of higher education institutions should determine to what extent their institutions can consider digital solutions to adapt to the changing nature of the sector “Education.” The activities supporting the transition of digital transformation processes by universities in Poland include (own elaboration based on [Mazurek, 2019]):

- supporting universities in increasing digital competencies, centralizing information systems, and making decisions based on the most accurate data sets and database analytics,
- supporting scientists, teachers, and university administration employees in rapid, even radical improvement of their digital competencies,
- preparing qualified managerial staff who, understanding contemporary changes, will be able to prepare universities for the challenges of digital transformation,
- changing the work culture at universities - from functional hierarchies and a "silo" approach to multi-task teams working on a project basis,
- increasing digital and analytical competencies of the administrative apparatus, as well as educators and scientists,
- a conscious approach to data collected at the university, including IT systems,
- continuous improvement of communication and research on the needs / motivation and commitment of employees and students,
- changes in the methods of education, the use of interactive tools, and alternative ways of working with the student.

Research, administrative, and especially teaching staff of universities should be ready to adopt new digital solutions and ensure the appropriate use of technology in everyday work being aware that it is currently one of the keyways to increase the attractiveness of studying and improving positioning on the international science market.

Employees should be provided with time and properly prepared for changes related to digitization / virtualization. It is very important to ensure a sense of security through effective communication at various stages of change.

It seems that the maximum degree of digitization is justified in extreme conditions, such as e.g., war or pandemic, while under normal/everyday conditions, the decision on the degree and areas of digitization of a particular university should be made after thorough analysis and considering many factors, both financial, organizational, and social, which include:

- financial opportunities related to the acquisition and maintenance of appropriate infrastructure and technical support,
- organizational possibilities (also the time horizon needed to implement changes (preparation of infrastructure, conducting the necessary training, preparation of new curricula and syllabuses considering new forms of conducting classes / selected classes),
- the specificity of the university,
- specificity of fields of study / specialization - different fields of study may have different needs in the digitization of didactic processes (e.g., classes in medical, chemical, etc. seem to be the least susceptible to digitization),
- expectations and concerns of both candidates for studies and the staff themselves,
- possible negative psychological impact - breaking ties or preventing the creation of ties between various participants of processes carried out at the university; in extreme cases - depression or other mental problems related to not finding oneself in new work / study conditions.

Also, too much offer for teachers and students who cannot (within a limited time) effectively use modern ICT tools in education, may constitute another challenge for universities.

IV. The use of Digitization in Didactic Processes

The progressive digitization in society, the dynamically changing market of educational needs with new priorities in the global educational market, and the development of the knowledge-based economy are the processes that determine a new approach to the learner in the broadly understood educational process. Nowadays, the digitization of education is not only an idea or a challenge but a necessity.

The ever-increasing use of the latest technologies in education is mainly explained by the new opportunities have been created and are still creating. The use of ICT in teaching processes intensifies and accelerates the process of absorption and assimilation of knowledge, and thus increases the effectiveness of teaching. Moreover, their use has a positive effect on the creation and development of key competencies that are often missing by university graduates, such as analytical and critical thinking, problem-solving, or the ability to cooperate and share knowledge [8]. The technologies used, imposing specific forms of organization of information transfer, also bring about cultural changes, change the scope of knowledge transfer, re-evaluating the view of the education process, and at the same time making it more attractive.

Many discussions on changes in the way of teaching come down to answering the question: Will ICT replace the traditional teaching model? Today it can only be said that the changes to come will undoubtedly result in the inclusion of new learning and teaching paradigms. The initiative to create new solutions will be transferred to the learners themselves, stimulating their motivation and activation. These activities will contribute to the development of learning, but only if the creativity of teachers, high-quality digital resources and educational applications are combined [9]. Such a compilation (along with the didactic process carried out) will allow to find a common space for communication, understanding, education and knowledge creation and its implementation in your own development area.

Modern technologies not only support the didactic process but also change the way universities operate. The use of modern ICT tools in the process of educating students undoubtedly determines not only a change in its quality but also an increase in the competitiveness of universities.

When deciding to start studying at a university, future students pay special attention to the way classes are conducted. Traditional forms of knowledge transfer are becoming schematic and increasingly unattractive. Therefore, it is important to constantly modify the teaching process at universities. Only interesting forms of knowledge transfer have a significant impact on the mental functions of the listener, significantly affect his interest and, above all, the development of competencies. The use of modern ITC tools in teaching students particularly relates to the development of competencies [10]:

- cognitive:

• digital and analytical competencies related to thecreation/usee of information, media, technological efficiency,

• communication skills, the ability to solve problems independently, inventive thinking, creativity,

- action-oriented:

• independence, flexibility, time management skills,

• productivity and use of digital resources,

- social:

• cooperation involving interaction based on cooperation, emotional management,

• teamwork.

The use of this type of solutions in teaching students helps to get to know reality more thoroughly, transform it rationally and stimulate students' creative activity. Summing up it should also be noted that the progressive digitization of education, especially during and after the COVID-19 pandemic, causes an increasing revival of educational communities around universities and their tasks. One of the reasons is the significant increase in competition among providers of digital teaching systems, programs, and tools.

V. Conclusion

The authors wanted, by presenting both the potential and barriers related to the digitization of universities, to contribute to the discussion whether digitization of universities is a necessity /future of universities. Recognizing the digitization of higher education as an inevitable direction of change, they tried to emphasize in which areas and under what conditions it should be implemented to negatively affect the entities participating in it (primarily staff and students) as little as possible.

Educational processes, the manifestation of which are, inter alia, classes with students are difficult to standardize in hundred percent, hence they are problematic and demanding to digitize. Their nature, however, does not exclude partial digitization to achieve the benefits described in this article.

In addition to indicating the potential and unquestionable benefits of developing virtualization and digitalization of higher education, the article also pointed a few important obstacles and doubts that universities must consider not to lose confidence and at the same time to keep up with the dynamically changing needs of the environment / market. These are both individual barriers (related to the attitude / motivation of staff and students to use different ICT tools and platforms and being literally "connected" to a large virtual system of the university with all its repercussions), and organizational, financial and technological ones, related to the need to properly plan this type of changes (e.g., creating appropriate infrastructure, providing technical support, and preparation and planning of information campaigns and trainings in various groups of interested parties).

However, it should be remembered that the teacher-student relationship plays a key role in the teaching process. One can try to dehumanize it, but before that happens, the question should be asked - if and who wants it. Hence, according to the authors, it is worth investigating the motives for taking actions related to the virtualization and digitization of universities; ask questions for whom and for what purpose it will be used; or the thesis that it is impossible to avoid the process of digitization of didactics is not only a justification for the pursuit of total control over what is happening in the classroom without leaving any margin for spontaneity, creativity, and sometimes creative improvisation adapted to the needs of the moment - overestimating codification to the detriment for interpersonal relationships and face to face communication. Prof. T. Pietrzykowski very aptly sums up the discussion on the digitization of universities: „The future is not a virtual university, that is, an excellent university such as Cambridge or Harvard, which attracts people to several years of online studies. This path will not work, because such education has one fundamental drawback: it does not equip students with the social competencies that are so needed today. That is why let us digitize universities, but with a view to complementary online contact education" [Tytuła 2020].

References

[1] P. Petrov, M. Radev, G. Dimitrov, G., & D. Simeonidis, “Infrastructure Capacity Planning in Digitalization of Educational Services,” 2022, International Journal of Emerging Technologies in Learning (IJET), 17(3), 299–306

[2] M. Cuypers, “Internationalization through web-based learning? An assessment of the virtualization of German universities,” 2012, retrieved on June 2 from: http://www.issbs.si/press/ISBN/978-961-6813-10-5/papers/ML12_074.pdf

[3] T. Pfeffer, “Virtualization of Research Universities: Raising the Right Questions to Address Key Functions of the Institution,” 2003, retrieved on June 2 from: https://escholarship.org/uc/item/6bv9c4qw

[4] K. O. Prosyukova, F.F. Shigapova, “Virtualization and digitalization in higher education,” 2020, ACM International Conference Proceeding Series, 1–3. https://doi.org/10.1145/3388984.3389063

[5] M. Tytuła, „Digitalizacja internacjonalizacji – tak, ale jak?” 2020, retrieved on June 2 from: https://perspektywy.pl/portal/index.php?option=com_content&view=article&id=7130:digitalizacja-tak-ale-jak&catid=24&Itemid=119

[6] A. Binsztok, B. Butryn, K. Hołowińska, L.M. Owoc & M. Sobińska, „Business computer simulation supporting competencies. Potential areas of application and barriers,” accepted for publication, KES 2022 conference, to be published.

[7] G. Mazurek, „Transformacja cyfrowa perspektywa instytucji szkolnictwa wyższego,” in: J. Woźnicki (ed.), Transformacja Akademickiego Szkolnictwa Wyższego w Polsce w okresie 1989–2019, 2019.

[8] P. Klimas, „Stosowanie mediów cyfrowych w edukacji wyższej – konfrontacja opinii nauczycieli akademickich i studentów” Wydawnictwo Naukowe UAM, Poznań, Studia Edukacyjne, nr 49, 2018, 269-280.

[9] I. Rudnicka, „W stronę nauczania mobilnego – prezentacja i wizualizacja treści wsparciem dla ucznia cyfrowej szkoły”, Ogólnopolskie Sympozjum Naukowe „Człowiek - Media - Edukacja” Uniwersytet Pedagogiczny w Krakowie 27-28 września 2013, retrieved on June 2 from: https://ktime.up.krakow.pl/symp2013/referaty_2013_10/rudnicka.pdf.

[10] A. Gaweł, „Rozwój kompetencji przedsiębiorczych dzięki nauczaniu z wykorzystaniem wirtualnych gier strategicznych”, Edukacja Ekonomistów i Menedżerów 2 (48), 2018.

Position Papers of the of the 17th Conference on Computer Science and Intelligence Systems pp. 113–120 DOI: 10.15439/2022F271 ISSN 2300-5963 ACSIS, Vol. 31

Generative Adversarial Networks for students' structure prediction. Preliminary research

Agata Kozina, Krzysztof Nowosielski, Zdzisław Kes,
Olena Sidor, Marcin Hernes, Paweł Golec
Wroclaw University of Economics and Business
ul. Komandorska 118/120, 53-345 Wrocław, Poland
Email: {agata.kozina, krzysztof.nowosielski, zdzisław.kes,
olena.sidor, marcin.hernes, pawel.golec}@ue.wroc.pl

Korlan Zhanat
Almaty Management University
Rozybakiev Street 227, Almaty 050060, Kazachstan
Email: k.zhanatovna@gmail.com

Abstract— **The effectiveness of the university's functioning and its organizational culture can be improved thanks to the use of machine learning. At Universities, the context of student anticipation is very important from the point of view of the fundamental planning and control functions associated with this specific form of management. The purpose of this study is to present the results of an experiment involving the prediction of student structure (attributes of students and their activities) based on the use of a machine learning solution and comparing them against real data obtained from a registry system of a European public institution of higher education in economic sciences. At universities, there is a clear need to support various components of system management. The experiments revealed that - for 11 out of the 48 examined datasets - the Percentage Similarity Index was in excess of 75% but was decidedly lower for the remaining sets (with 18 sets assessed below the margin of 50%).**

I. Introduction

A identifies a set of properties deemed important in the context of enrolment management are consisted from [1], [2]:

- factors that induce and incite university enrolment,
- proper understanding of reasons behind dropouts as well as incentives for persistence in students,
- forms of financing employed by students to cover the cost of their education,
- strategic planning of tasks related to the university's present and future financing needs,
- integration between enrolment management and retention management tasks.

In general, the principal function of enrolment management is to provide effective control over student characteristics and student population size.

As observed by Dixon (1995), enrolment management may be designed in pursuit of the following four objectives: (1) clear definition and propagation of institutional goals, (2) ensuring stakeholders' full support for marketing plans and activities made in relation to institutional goals, (3) making strategic decisions on the role and volume of financial aid required to reach and retain the desired size of the student population, and (4) making significant commitments for the realisation of the above. Enrolment management exerts a significant impact upon the structure of the student population and, consequently, the structure of university revenues, as forms of service and curricula are directly manifested in tuition costs, and thus determine the financial standing of the institution [3]. According to a European University Association report, several distinct trends can be observed in the development of public financing and student enrolment in the years 2008-2016 in Europe [4], [5]. In Poland, intensive efforts are underway at present to increase public support for higher learning to negate the effects of brain drain and the gradual decrease of the student population. The above aspects clearly emphasise the need for a more proactive design of the institutional enrolment policy as an essential determinant of future tuition revenues, resource allocation for subsequent academic years, and the creation of marketing plans, especially ensuring their adjustment to specific segments of the university's offer.

In the context of increased competition among the national universities, the strengthening of the influence of global educational processes on the domestic higher education, the need to change the management component of the system becomes obvious. In their development, universities face a large number of challenges, such as: the development of technologies, the commercialization of activities, the increase in the amount of information, the changing requirements of employers for graduates as potential employees. One of the research problems of higher education management is the prediction of students structure. Many higher-education institutions are now using data and analytics as an integral part of their processes. Whether the goal is to identify and better support pain points in the student journey, more efficiently allocate resources, or improve student and faculty experience, institutions are seeing the benefits of data-backed solutions.

The aim of the paper is to develop the method for students' structure prediction using machine learning.

II. BACKGROUND

Corporate culture in the organization arises regardless of whether it is planned from above or not. It exists as a given in any organization, even in a newly created company, it is created by the employees themselves. At the same time, it can either help in achieving the goals of the company, or slow down this process. Corporate culture determines how employees approach problem solving, interact with each other, behave in conflict situations, serve customers, deal with suppliers, and how they generally carry out their activities [6].

To organize effective work, it is necessary to use all available management methods. These methods, according to the authors of the book "Methods of personnel management" are divided into economic, administrative-legal and socio-psychological [7].

Corporate governance always relies on both formal and informal structures. The formal structure is based on the norms that are mandatory for the organization's personnel: hierarchy of subordination, unity of command, sanctions, coercion. The informal structure is based on norms associated with values: sympathy, authority, collegiality, initiative. It is obvious that the use of one formal management leads to rigidity, lack of flexibility in the organization, hinders the development of initiative, which hinders the further development of the organization and leads to loss in competition, while the predominance of the informal structure will lead to chaos, loss of control over the entire hierarchical structure. chain. Therefore, a necessary condition for successful management is the fulfillment of two requirements:

- decentralization of powers and responsibilities within the company to certain limits;
- formation of a single team of employees of the organization.

Currently, universities are mainly characterized by "Club culture", which allows them to work efficiently and smoothly [1].

There are a number of phases of employee interaction with the corporate culture of the university [8]:

1. At the first stage (orientation phase), the employee gets acquainted with the mission, values, symbols of the university, using Internet sites and other information materials;
2. At the second stage (adaptation phase), adaptation to the corporate culture of the university takes place;
3. At the stage of interaction, immersion into the value system of the university takes place, a wide range of communicative interactions with various groups is carried out;
4. The phase of integration involves the value unity of the university with the employee as a bearer of corporate culture.

Those at the forefront of this trend are focusing on harnessing analytics to increase program personalization and flexibility, as well as to improve retention by identifying students at risk of dropping out and reaching out proactively with tailored interventions. Indeed, data science and machine learning may unlock significant value for universities by ensuring resources are targeted toward the highest-impact opportunities to improve access for more students, as well as student engagement and satisfaction [10].

Yet higher education is still in the early stages of data capability building. With universities facing many challenges (such as financial pressures, the demographic cliff, and an uptick in student mental-health issues) and a variety of opportunities (including reaching adult learners and scaling online learning), expanding use of advanced analytics and machine learning may prove beneficial.

Below, we share some of the most promising use cases for advanced analytics in higher education to show how universities are capitalizing on those opportunities to overcome current challenges, both enabling access for many more students and improving the student experience[11].

Data science and machine learning may unlock significant value for universities by ensuring resources are targeted toward the highest-impact opportunities to improve access for more students, as well as student engagement and satisfaction.

Advanced analytics—which uses the power of algorithms such as gradient boosting and random forest—may also help institutions address inadvertent biases in their existing methods of identifying at-risk students and proactively design tailored interventions to mitigate the majority of identified risks. For instance, institutions using linear, rule-based approaches look at indicators such as low grades and poor attendance to identify students at risk of dropping out; institutions then reach out to these students and launch initiatives to better support them. While such initiatives may be of use, they often are implemented too late and only target a subset of the at-risk population [4]. This approach could be a good makeshift solution for two problems facing student success leaders at universities. First, there are too many variables that could be analyzed to indicate risk of attrition (such as academic, financial, and mental health factors, and sense of belonging on campus). Second, while it's easy to identify notable variance on any one or two variables, it is challenging to identify nominal variance on multiple variables.

III. MATERIALS AND METHOD

A. Input data characteristic

As already established, input data includes two groups of variables:

- dependent variables (set of 9883 records, with each record described by 15 attributes) obtained from the registry system of the examined university for the years 2016-2020;
- independent variables (3 attributes) obtained from the national statistical records published by the Central Statistical Office

The variables were coded as follows:

- X1 - work_name – entities employing the candidates were divided according to the type of business;
- X2 - code – fields of studies were grouped by subject;
- X3 - work_city – places of student residence were coded by their physical distance from the university (in km);
- X4 - nationality – the nationality of students;
- X5 - gender – gender of students was coded as follows: 0 for males, 1 for females;
- X6 - status – codes of student status;
- X7 - finished_university – the enrolment data provides details of each candidate's previous education. The recorded institutions of higher learning were assigned codes from 0 to 364.
- X8 - work_years – work experience of candidates registered in the database (in years of service).

B. The model's learning method

GANs are a relatively new method in the field of machine learning. These networks, which were introduced in 2014 by Ian Goodfellow and his collaborators, are designed to create new data that in some form mimics the statistical properties of a given set of training data. Given a target dataset, such as celebrity faces or categories from the ImageNet dataset, a GAN can be trained that generates new, unseen data that (ideally) fit comfortably and indistinguishably in the dataset. Since the introduction of GANs, several variations of the architecture and many theories to help train these inherently unstable networks have been developed[11].

In general, GANs are composed of generator and discriminator neural networks (Figure 1), which, for image data, are typically convolutional networks.

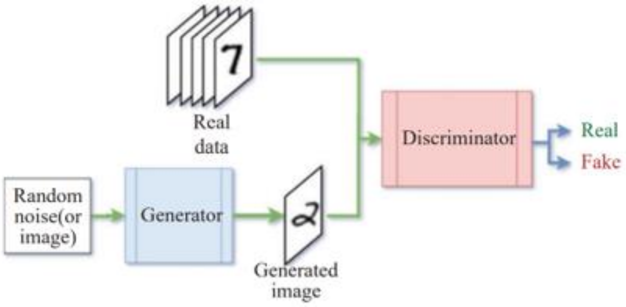


Fig. 1 A diagram of a generic generative adversarial network. The network shown here is designed to produce new images of handwritten MNIST digits. The generator converts random noise into images that attempt to match the data from the target dataset. The discriminator distinguisges between real and generated data

Training is accomplished by repeatedly presenting the networks with data from a target dataset. The generator is tasked with learning to convert random n-dimensional vectors to data matching the dataset. The discriminator, in turn, is tasked with distinguishing between data from the dataset and the generator's output. In a descriptive analogy offered by Goodfellow et al., the generator can be likened to an art forger, the goal of which is to create undetectable forgeries of the world's great artists. The discriminator plays the role of a detective, trying to discover which pieces are real and which are fakes. The loss function for a GAN is given by

$$\begin{aligned}\min\nolimits_G \max\nolimits_D L(D,G) &= \mathbb{E}_{x\sim p_r(x)}[\log(D(x))] + \mathbb{E}_{z\sim p_z(z)}[\log(1-D(G(z)))] \\ &= \mathbb{E}_{x\sim p_r(x)}[\log(D(x))] + \mathbb{E}_{x\sim p_g(x)}[\log(1-D(x))]\end{aligned} \tag{1}$$

where $G(z)$is the output of generator network, $D(x)$ is the output the discriminator network, z is a multidimensional random input to the generator, p_z is the distribution of z (usually uniform), and $p_g(x)$ and $p_r(x)$ are the probability manifold distributions for the generated data and the target dataset, respectively. Via backpropagation, these objectives direct the generator to create data that fits well with the dataset, while simultaneously increasing the distinguishing power of the discriminator. It can be shown[11] that, for a GAN with sufficient capacity, this training objective minimizes the Kullback-Leibler divergence between $p_g(x)$ and $p_r(x)$. This divergence metric describes how similar two probability distributions are, with low values denoting greater similarity. In other words, training a GAN creates a generator that is able to mimic the distribution of data in the given dataset at some level.

Models of the generator and the discriminator are presented below. The Sequential model utilises the following layers: Dense, LeakyReLU oraz BatchNormalization. All the layers and the model itself are derived from the Keras library.

Tables VI and VII present the structure of both models. The generator model includes five layers of 'Dense', two layers of batch normalisation, and two functions Leaky ReLU, serving as activation functions. The discriminator model employs seven layers of 'Dense', four functions Leaky ReLU as activation functions for the neurons defined above. The 'Output Shape' column reports a number of nodes for each layer. The loss function used in the discriminator model was developed on the basis of the Binary Cross Entropy function defined as follows (Eq. 1):

$$Loss(y_i, z_i) = \begin{cases} z_i - z_i y_i + \log(1 + e^{-z_i}) & if\ z_i \geq 0 \\ -z_i y_i + \log(1 + e^{z_i}) & if\ z_i < 0 \end{cases} \quad (2)$$

Both models (the generator and the discriminator) are fed with discriminator responses (argument z$_i$):

- The generator's loss function is the Binary Cross Entropy function with values
  $$y_i = 1, for\ \forall_{i=0}^{N} z_i$$
- Discriminator: sum of Binary Cross Entropy functions with values
  $$y_i = \begin{cases} 1, & for\ real\ z \\ 0, & for\ artificial\ z \end{cases}$$

Both models utilise the 'Adam' algorithm with a learning step: 0.0001. This step value has already been employed in GANs procedures for TensorFlow. At the same time, as evidenced by research presented in [12], the value yields much better results compared to other algorithms, offering the added benefit of facile and simple implementation in Keras.

For the entire duration of the learning process, examples were fed randomly. The network gained knowledge of the student patterns based on the entire set of input data. The training procedure was set at 55 000 iterations. One epoch was represented by one packet of data holding information on 16 students. Figure 2 provides a plot of the training history.

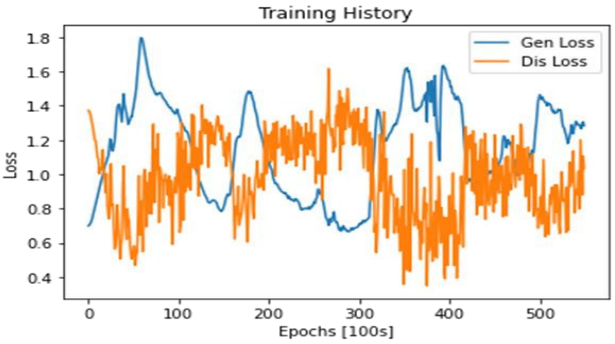


Fig. 2 Training history

As evidenced by the above, the discriminator was able to recognise between fake and real data at a relatively early stage of the training process. This was accompanied by deterioration in the quality of the generator's output over time. This phenomenon can be explained by differences in the number of layers. As the generator utilised fewer layers than the discriminator, its training processes were more immediate. However, the discriminator was, at the same time, more effective in its long-term predictions, owing to the benefit of more layers. The continual learning phenomenon was established.

Table VI.
Generator. Model: "sequential"

Layer (type)	Output Shape	Param #
dense (Dense)	(None, 4)	20
batch_normalization_3	(Batch (None, 4)	16
leaky_re_lu_9 (LeakyReLU)	(None, 4)	0
dense_1 (Dense)	(None, 5)	25
dense_2 (Dense)	(None, 6)	36
dense_3 (Dense)	(None, 7)	49
dense_4 (Dense)	(None, 7)	56
batch_normalization_1	(Batch (None, 7)	28
leaky_re_lu_1 (LeakyReLU)	(None, 7)	0
dense_5 (Dense)	(None, 8)	64
Total params: 294 Trainable params: 272 Non-trainable params: 22		

Table VII.
Discriminator. Model: "sequential_1"

Layer (type)	Output Shape	Param #
dense_6 (Dense)	(None, 7)	63
dense_7 (Dense)	(None, 6)	48
leaky_re_lu_2 (LeakyReLU)	(None, 6)	0
dense_8 (Dense)	(None, 5)	35
leaky_re_lu_3 (LeakyReLU)	(None, 5)	0
dense_9 (Dense)	(None, 4)	24
leaky_re_lu_4 (LeakyReLU)	(None, 4)	0
dense_10 (Dense)	(None, 3)	15
leaky_re_lu_5 (LeakyReLU)	(None, 3)	0
dense_11 (Dense)	(None, 2)	8
dense_12 (Dense)	(None, 1)	3
Total params: 196 Trainable params: 196 Non-trainable params: 0		

C. Methods of output data verification

The Percentage Similarity Index (PSI) was adopted to verify the established similarities between structures of individual variables and those generated by GANs. The index was calculated for equinumerous sets of structural indices based on formula (Eg. 2).

$$PSI = \sum_{k=1}^{n} min(I_{1k}; I_{2k}) \qquad (2)$$

where:

- PSI - percentage similarity index,
- I_{1k} - percentage share of k-th component in the structure of set 1,
- I_{2k} - percentage share of k-th component of the structure of set 2,
- n - number of elements in set 1 (both sets need to be equinumerous).

The following similarity ranges were defined for the evaluation of the sets:

- 100% - 90% - sets are similar,
- 90% - 75% - sets are moderately similar,
- 75% - 50% - similarity between sets is marginal,
- 50% - 0% - sets are not similar

IV. Research results and discussion

A. Output data

The results obtained from the trained generator in the course of the experiment, complete with student characteristics, statistical properties and examples derived from the output data set, are presented below. Table VIII presents a segment of output data generated by GANs for the year 2021.

Each column of the generated output corresponds to specific information items stored in the university database of student records. Parts of the output data were rounded off, as dictated by the specificity of information stored therein. An example of such procedure is the 'Status' column, with domain defined by $x\in\langle 0;4\rangle \wedge x\in Z$.

B. Output data verification

As suggested by the statistical properties of data, the generated output records are well contained in the brackets defined by the real data. It was assumed that the structure of output data generated by the GEN network for the years 2016-2020 should take up values similar to those of the real records stored for the period. The PSI was used to verify the similarity between the structures of individual variables and the output data generated by the GEN network. Results of the output data verification procedure are provided in Table IX. The PSI exceeded 75% for eleven cases among all tested variables, which suggests their similar or moderately similar character. Thus, it may be concluded that in those cases, the training turned out to be consistent with these segments of data. The best results were obtained for variable X4, nationality – PSI for all tested years was over 90%, and the overall value was 93.5%. Quite satisfactory levels of PSI, between 60 and 90%, were received for X3, city, X5, gender, and X7, work experience, overall PSI reached respectively 73.5%, 70.4%, and 67.2%. The worst fit, overall PSI 30.9%, was found for X8, university. For each variable, histograms were produced to observe the similarities between the representations of both datasets. Figures below present percentage shares in the categories represented in the real dataset and the set generated by the GEN network. Fig. 3 illustrates the structure of the 'employer' variable (X1) for real and generated data for the tested years. The overall PSI index calculated for variable X1 was at 44.5%, suggesting a dissimilarity between the structures generated by GENs and those of the real data. The structure of the 'study field' variable (X2) for real and generated data for 2016-2020 is shown in Fig. 4. The PSI index calculated for variable X2 was 52.3%, suggesting a

TABLE VIII.
CHARACTERISTICS OF REAL DATA

Parameter	Employer	Study field	City	Nationality	Gender	Status	Work experience	Finished University
mean	2.456	4.973	53.954	0.020	0.674	1.051	3.386	44.747
std	2.261	2.857	123.136	0.331	0.469	0.843	6.215	71.159
min	0.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000
25%	0.000	3.000	0.000	0.000	0.000	0.000	0.000	2.000
50%	3.000	4.000	0.000	0.000	1.000	1.000	0.000	3.000
75%	4.000	6.000	70.200	0.000	1.000	1.000	4.000	78.000
max	6.000	12.000	3 349.000	12.000	1.000	4.000	42.000	364.000

TABLE IX.
PERCENTAGE SIMILARITY INDEX (PSI) FOR TESTED VARIABLES

Variable		Years					All Years
		2016	2017	2018	2019	2020	
Employer	X1	46.2%	47.5%	41.3%	44.7%	37.9%	44.5%
Study field	X2	49.7%	49.3%	51.4%	52.4%	5.8%	52.3%
City	X3	68.9%	74.0%	72.8%	78.5%	71.9%	73.5%
Nationality	X4	90.1%	90.9%	95.9%	95.3%	99.5%	93.5%
Gender	X5	68.7%	66.9%	71.2%	71.3%	87.7%	70.4%
Status	X6	43.6%	90.3%	71.0%	15.6%	0.0%	54.0%
Work exp.	X7	77.1%	82.1%	59.4%	53.8%	50.7%	67.2%
University	X8	25.6%	33.3%	30.6%	31.8%	6.7%	30.9%

medium similarity between the structures generated by GENs and those of the real data. Data stored in the real records of the university identify management as the most attractive field of study. In contrast, output data generated by the network reported Audit and financial control as the dominant area.

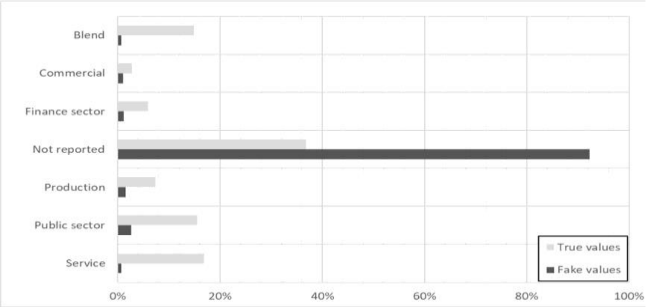


Fig. 3 Histogram of the 'employer' variable (X1)

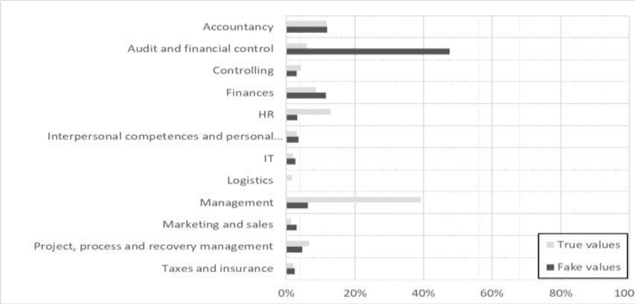


Fig. 4 Histogram of the 'study field' variable (X2)

Fig. 5 presents the structure of the 'city' variable (X3) for real and generated data.

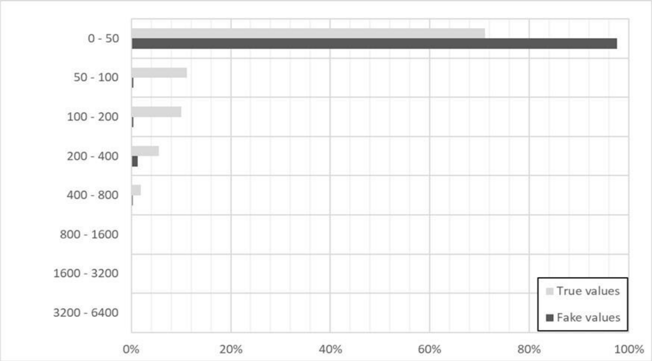


Fig. 5 Histogram of the 'city' variable (X3)

The overall PSI index calculated for variable X3 was 73.5%, suggesting a significant similarity between the structures generated by GENs and those of the real data. Real data shows that Wroclaw (including the city outskirts) is the place of residence for the overwhelming majority of the student population. The structure of GENs output data suggests that ca. 71% of students commute over a distance of 0 to 50 km, which is compatible with real data.

The structure of the 'nationality' variable (X4) for real and generated data is illustrated in Fig. 6. The overall PSI index calculated for variable X4 was 93.5%, suggesting a strong similarity between the structures generated by GENs and those of the real data. The real data shows that more than 99% of postgraduate students come from Poland. The structure of GENs output data suggests that ca. 7% represent a foreign nationality.

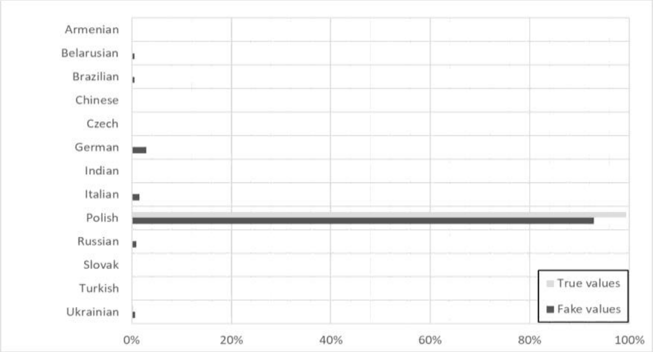


Fig. 6 Histogram of the 'nationality' variable (X4)

Fig. 7 presents the structure of the 'gender' variable (X5) for real and generated data for the tested years. The overall PSI index calculated for variable X5 was 70.4%, suggesting a significant similarity between the structures generated by GENs and those of the real data. The actual data shows that female students account for two-thirds of the student population, while the data predicted by GENs shows complete female dominance.

The structure of the 'status' variable (X6) for real and generated data is shown in Fig. 8. The PSI index calculated for variable X6 was 54.0%, suggesting a medium similarity between the structures generated by GENs and those of the real data. Generated data indicate that most of the students are promoted, while the real data show that many students are still studying, have been deleted or resigned.

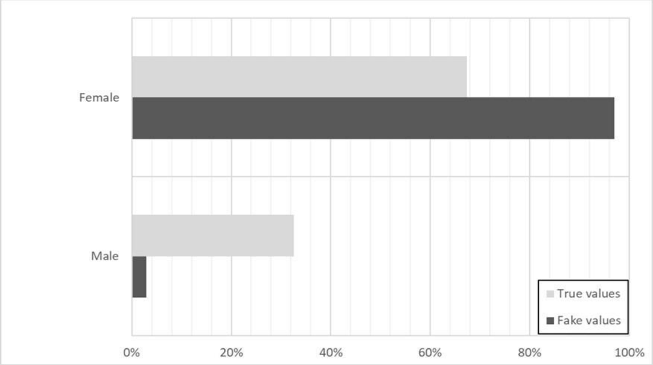


Fig. 7 Histogram of the 'gender' variable (X5)

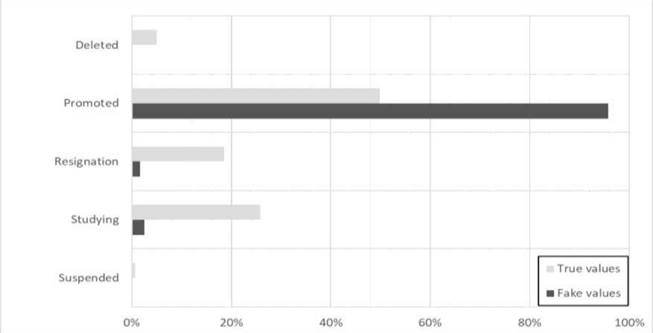


Fig. 8 Histogram of the 'status' variable (X6)

Fig. 9 presents the structure of the 'work experience' variable (X7) for real and generated data for the tested years. The PSI index calculated for variable X8 was at 67.2%, suggesting a moderate similarity between the structures generated by GENs and those of the real data. Generated data show that most of the students have worked for two years or less, while the real data show also that the work experience in many cases is longer.

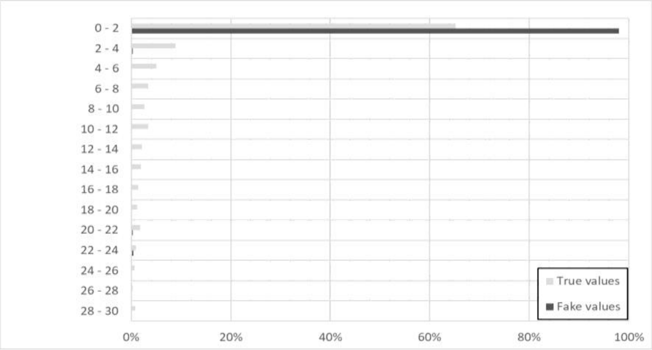


Fig. 9 Histogram of the 'work experience' variable (X7)

The structure of the 'finished university' variable (X8) for real and generated data for the years 2016-2020 is illustrated in Fig. 10. The PSI index calculated for variable X8 was at 30.9%, suggesting a dissimilarity between the structures generated by GENs and those of the real data.

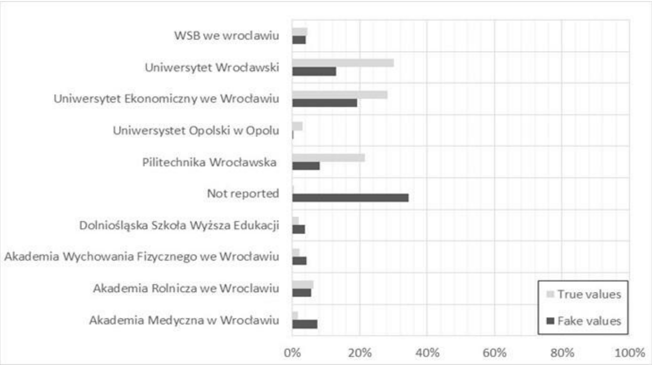


Fig. 10 Histogram of the 'finished university' variable (X8)

Based on analytical evaluations, it may be concluded that the topic of student structure prediction (SSP) is not adequately represented in the professional literature. At the same time, the studied concept serves important practical purposes and presents a major challenge for the managerial cadres of public institutions of higher education. Another important issue of this paper is the use of GANs in predictions. The use of such tools is presented in the literature. However, from the viewpoint of this research and the instrumental utilisation of this technique in SSP, this particular segment of knowledge should not be treated as a point of reference in our discussion. There remains a key area for the purpose of the work, i.e. the use of GANs in the SSP. In this context, scarcity of reference material can be observed, similar to that of the SSP. The available literature is limited to the prediction of the number of students or students' performance. Naturally, these aspects are in close association with SSP, but they are too far detached from the nature of SSP to be of any rational significance in the studied context. Their practical benefits may be important from a broader perspective of using machine learning solutions as a form of management support in university administration. Because of the observed scarcity of reference material related to the studied context, the authors of this study propose to treat the presented research results as a contribution to the discussion on the use of GANs in the SSP.

V. CONCLUSION AND IMPLICATIONS

Working on a copious set of factors of potential impact upon the student structure prediction presented in this paper, the authors examined the perspective of applying 'intelligent solution' methods for the task performed based on Generative Adversarial Networks. The research was conducted on a dataset of records describing the real population of students of postgraduate studies over a period of 5 academic years, between 2016 and 2020. Individual properties and attributes of students were coded. The dataset was supplemented by a number of indices describing the general economic condition of the region proper for the studied university and the timeframe under study. The final design included 12 dependent variables and three independent variables to give a total of 15 variables. The experiment made use of artificial intelligence networks, specifically the GANs networks. The network was presented with the tasks or reproducing the structure of students to produce output adequately comparable with real data recorded for previous years. The Percentage Similarity Index (PSI) was calculated for each variable to illustrate the similarity between their real structure and that produced by GANs. The experiments revealed that – for 11 out of the 48 examined datasets – the PSI index was in excess of 75% but was decidedly lower for the remaining sets (with 18 sets assessed below the margin of 50%). This should be interpreted as evidence that only parts of data generated by GANs sufficiently reflect the real data. Additional tests may be required to provide grounds for more reliable predictions of student structure, including those involving different sets of independent variables. The need for extension of the set of variables is fairly evident. More effort should be placed to verify their information potential and activate a learning mechanism after verifying or exchanging variables. Other methods for selecting variables should also be examined, as the present set was established based on the expert method. Further research directions may involve the development of methods based on other neural network architectures (such as Recurrent Neural Networks, Convolutional Neural Networks) to predict postgraduate students' structure. The method applied by authors may not be an ideal solution to the problem at hand. However, since the attempt proved partially effective, the results are of scientific value and may serve as the basis for further examination of the SSP concept.

FUNDING

The project is financed by the Ministry of Science and Higher Education in Poland under the programme "Regional Initiative of Excellence" 2019 - 2022 project number 015/RID/2018/19 total funding amount 10 721 040,00 PLN

References

[1] D. Hossler and B. Bontrager, *Handbook of strategic enrollment management*. San Francisco, CA: Jossey-Bass, A Wiley Brand, 2014.

[2] M. J. Denniss, ‘Anticipatory Enrollment Management: Another Level of Enrollment Management’, vol. 88, no. 1, pp. 10–16, 2012.

[3] D. Trusheim and C. Rylee, ‘Predictive modeling: linking enrollment and budgeting’, *Planning for Higher Education*, vol. 40, 2011.

[4] D. M. West, ‘Big data for education: Data mining, data analytics, and web dashboards’, *Brookings*, 2012. https://www.brookings.edu/research/big-data-for-education-data-mining-data-analytics-and-web-dashboards/

[5] X. Shacklock, ‘From bricks to clicks: the potential of data and analytics in higher education’, *VOCEDplus*, 2016. http://hdl.voced.edu.au/10707/411226

[6] E. N. Saribekyan, ‘Organizational culture and organizational culture’, *Culture: management, economics, law*, no. 4, p. 37, 2014.

[7] V. N. Fedoseev and S. N. Kapustin, ‘Methods of personnel management’, in *Analysis of the crop industry*, Almaty, 2014.

[8] A. V. Shelyakina, ‘Corporate culture of the organization’, *Young scientist*, no. 14, pp. 206–209, 2018.

[9] M. V. Shumeiko, ‘Typology of corporate culture’, *Cyberleninka*, p. 8.

[10] J.A. Gray and M. DiLoreto. "The effects of student engagement, student satisfaction, and perceived learning in online learning environments." *International Journal of Educational Leadership* Preparation 11.1 (2016): n1.

[11] S. Z. Gurbuz, Ed., ‘Machine learning techniques for SAR data augmentation’, in *Deep Neural Network Design for Radar Applications*, Institution of Engineering and Technology, 2020, pp. 163–206. doi: 10.1049/SBRA529E_ch6.

[12] J. Brownlee, ‘Gentle Introduction to the Adam Optimization Algorithm for Deep Learning’, *Machine Learning Mastery*, 2017. https://machinelearningmastery.com/adam-optimization-algorithm-for-deep-learning/

Software, System and Service Engineering

THE S3E track emphasizes the issues relevant to developing and maintaining software systems that behave reliably, efficiently and effectively. This track investigates both established traditional approaches and modern emerging approaches to large software production and evolution.

For decades, it is still an open question in software industry, how to provide fast and effective software process and software services, and how to come to the software systems, embedded systems, autonomous systems, or cyber-physical systems that will address the open issue of supporting information management process in many, particularly complex organization systems. Even more, it is a hot issue how to provide a synergy between systems in common and software services as mandatory component of each modern organization, particularly in terms of IoT, Big Data, and Industry 4.0 paradigms.

In recent years, we are the witnesses of great movements in the area of software, system and service engineering (S3E). Such movements are both of technological and methodological nature. By this, today we have a huge selection of various technologies, tools, and methods in S3E as a discipline that helps in a support of the whole information life cycle in organization systems. Despite that, one of the hot issues in practice is still how to effectively develop and maintain complex systems from various aspects, particularly when software components are crucial for addressing declared system goals, and their successful operation. It seems that nowadays we have great theoretical potentials for application of new and more effective approaches in S3E. However, it is more likely that real deployment of such approaches in industry practice is far behind their theoretical potentials.

The main goal of Track 5 is to address open questions and real potentials for various applications of modern approaches and technologies in S3E so as to develop and implement effective software services in a support of information management and system engineering. We intend to address interdisciplinary character of a set of theories, methodologies, processes, architectures, and technologies in disciplines such as: Software Engineering Methods, Techniques, and Technologies, Cyber-Physical Systems, Lean and Agile Software Development, Design of Multimedia and Interaction Systems, Model Driven Approaches in System Development, Development of Effective Software Services and Intelligent Systems, as well as applications in various problem domains. We invite researchers from all over the world who will present their contributions, interdisciplinary approaches or case studies related to modern approaches in S3E. We express an interest in gathering scientists and practitioners interested in applying these disciplines in industry sector, as well as public and government sectors, such as healthcare, education, or security services. Experts from all sectors are welcomed.

TOPICS

Submissions to S3E are expected from, but not limited to the following topics:

- Advanced methodology approaches in S3E – new research and development issues
- Advanced S3E Process Models
- Applications of S3E in various problem domains – problems and lessons learned
- Applications of S3E in Lean Production and Lean Software Development
- Total Quality Management and Standardization for S3E
- Artificial Intelligence and Machine Learning methods in advancing S3E approaches
- S3E for Information and Business Intelligence Systems
- S3E for Embedded, Agent, Intelligent, Autonomous, and Cyber-Physical Systems
- S3E for Design of Multimedia and Interaction Systems
- S3E with User Experience and Interaction Design Methods
- S3E with Big Data and Data Science methods
- S3E with Blockchain and IoT Systems
- S3E for Cloud and Service-Oriented Systems
- S3E for Smart Data, Smart Products, and Smart Services World
- S3E in Digital Transformation
- Cyber-Physical Systems (9th Workshop IWCPS-9)
- Model Driven Approaches in System Development (7th Workshop MDASD'22)
- Software Engineering (42th IEEE Workshop SEW-42)

TRACK CHAIRS

- **Luković, Ivan,** Unniversity of Belgrade, Serbia
- **Kardas, ,** Geylani, Ege University International Computer Institute, Turkey

PROGRAM CHAIRS

- **Bowen, Jonathan,** Museophile Ltd., United Kingdom
- **Hinchey, Mike**(Lead Chair), Lero-the Irish Software Engineering Research Centre, Ireland
- **Szmuc, Tomasz,** AGH University of Science and Technology, Poland
- **Zalewski, Janusz,** Florida Gulf Coast University, United States
- **Seyed Hossein Haeri, ,** IOG and University of Bergen, Norway

PROGRAM COMMITTEE

- **Ahmad, Muhammad Ovais,** Karlstad University, Sweden
- **Challenger, Moharram,** University of Antwerp, Belgium
- **Dejanović, Igor,** Faculty of Technical Sciences, University of Novi Sad, Serbia
- **Derezinska, Anna,** Institute of Computer Science, Warsaw University of Technology, Poland
- **Dutta, Arpita,** National University of Singapore
- **Erata, Ferhat,** Yale University, USA
- **Escalona, M.J.,** University of Seville, Spain
- **Essebaa, Imane,** Faculté des Sciences et Techniques Mohammedia, Morocco
- **García-Mireles, Gabriel,** Universidad de Sonora, Mexico
- **Göknil, Arda,** SINTEF Digital, Norway
- **Hanslo, Ridewaan,** University of Pretoria, South Africa
- **Jarzebowicz, Aleksander,** Gdansk University of Technology, Poland
- **Kaloyanova, Kalinka,** University of Sofia, Bulgaria
- **Katic, Marija,** University of London, United Kingdom
- **Khlif, Wiem,** FSEGS, Tunisia
- **Kolukısa Tarhan, Ayça,** Hacettepe University, Turkey
- **Krdzavac, Nenad,** University of Belgrade, Serbia
- **Marcinkowski, Bartosz,** University of Gdansk, Poland
- **Milašinović, Boris,** Faculty of Electrical Engineering, University of Zagreb, Croatia
- **Milosavljevic, Gordana,** Faculty of Technical Sciences, University of Novi Sad, Serbia
- **Misra, Sanjay,** Ostfold University, Norway
- **Morales Trujillo, Miguel Ehécatl,** University of Canterbury, New Zealand
- **Ozkan, Necmettin,** Kuveyt Turk Participation Bank
- **Ozkaya, Mert,** Istanbul Kemerburgaz University, Turkey
- **Ristic, Sonja,** Faculty of Technical Sciences, University of Novi Sad, Serbia
- **Rossi, Bruno,** Masaryk University, Czech Republic
- **Sanden, Bo,** Colorado Technical University, USA
- **Sierra, Jose Luis,** Universidad Complutense de Madrid, Spain
- **Torrecilla-Salinas, Carlos,** IWT2
- **Varanda Pereira, Maria João,** Instituto Politécnico de Bragança, Portugal
- **Vescoukis, Vassilios,** National Technical University of Athens, Greece

DOI: 10.15439/2022F275
ISSN 2300-5963 ACSIS, Vol. 31

Small Footprint Embedded Systems Paradigm Based on a Novel and Scalable Implementation of FORTH

Bogusław Cyganek
AGH University of Science and Technology, Poland
Al. Mickiewicza 30, 30-059 Kraków, Poland
cyganek@agh.edu.pl

***Abstract*—This paper describes architecture of the novel implementation of the Forth interpreter-compiler. The architecture follows the object- and component-oriented design paradigms. The implementation is done with the modern C++ 20 language taking full advantage of such constructs as lambda functions, variadic templates, as well as the coroutines and concepts. The system is highly modular and easily scales for small footprint embedded systems. We propose to extend Forth with the coroutine words that allow for async operations and lightweight cooperative multi-threading. We show successful deployment of the proposed Forth implementation on three platforms, two PC frameworks running Linux and Windows, respectively, as well as on tiny embedded system NodeMCU v3 with the 32-bit RISC ESP8266 microprocessor and 32/80KB memory. The platform also has educational value, showing intrinsic operation of Forth and modern C++. Software is available free from the Internet.**

***Keywords*—: Forth, compiler-interpreter, multi-tasking, co-routines, co-operative systems, IoT**

I. Introduction

Forth is a computer language developed by Charles Moore in early 70s as a system to control the radio telescope when he worked in the National Radio Astronomy Observatory [11][13][19]. Its name was coined to commemorate the fourth generation of computers but since the file system restricted names to five letters only, Moore skipped the middle "U" and left Forth. The fascinating story of Forth is described in The Evolution of Forth [11], while a short biography of Charles Moore is in Wikipedia [13]. Forth has always been very outstanding, original and interesting computer language [1][10][12][14][15]. Although not in the mainstream, slightly forgotten today, we are deeply convinced it can still serve many purposes. This is especially true in the context of small embedded systems that need interactive features, such as ones for the Internet of Things (IoT), and also if Forth can be shown in the new light of a novel implementation in modern C++, as presented in this paper.

There are many free and commercial implementations of Forth, such as *Gforth*, which is a free GNU portable implementation of the ANS Forth standard for Linux/Unix, Windows, and other operating systems [20]. Another implementation is *Swift Forth*® by Forth Inc. [21]. On the other hand, a popular implementation with many follow ups is *jonesforth* project [22]. We just named few of the available projects, many more can be found online [19][23].

However, to the best of our knowledge, none of the above mentioned implementation uses modern C++, i.e. ver. 17 or 20 [5][24]. On the other hand, having a Forth implementation done with modern C++ allows to use the latest very efficient and productive features of C++, such as STL containers, variadic templates, on-time compilation, regular expressions, lambda functions, and coroutines. Especially the latter offers new ways of efficient implementation of the async IO operations, state machines, or lightweight multithreading, as will be discussed. Hence, the proposed implementation greatly reduces system complexity, at the same time allowing for scalable solutions. The complete Forth project presented in this paper, named *BCForth*, is available free from the Internet [16]. This also makes it a good teaching platform for the computer classes.

But most of all, what can be interesting in Forth when confronted with e.g. modern C++? The main difference is presence of the interpreter and compiler, at a relatively small footprint on the other hand. This means that, contrary to C++, which to add a new software component requires recompilation and rebuild, a Forth based system is very interactive and extensible. That is, the user can run the existing words but also can extend the system by his/her defined new words, which are immediately compiled and instantly become available for construction of next words, and so on. Not less important is the mentioned small footprint of Forth, which renders it useful for small embedded platforms, IoT, or even in the so called bare-metal systems. Hence, we can easily imagine a simple but smart sensor, which is run by Forth alone and allows communication, as well as extensions, in the run time.

The rest of the paper is organized as follows. Architecture of the Forth platform is presented in Section II. It is organized in four subsections: core architecture (II.A), key data structures (II.B), hierarchy of Forth words (II.C), and finally the system activity (II.D). The coroutine component – a proposed novel add-on to the Forth language – is dealt with in Section (III).

System deployment and experiments are presented in Section (IV). The paper ends with conclusions in Section (V).

II. Architecture of the Novel Forth Platform

The main purpose of the *BCForth*, is to provide a flexible implementation of Forth with the modern C++20, which can be easily ported to various embedded platforms endowed with the C++ compiler. *BCForth* contains also an extension in the form of the coroutines, as will be discussed. Contrary to some older implementations in assembly or C, modern C++ allows clear, understandable and extensible code. For instance, if necessary *BCForth* can be reduced of its components (e.g. it can run only with the interpreter), or it can be even ported to the older version e.g. C++ 11.

In this section we present basic assumptions behind the architecture of *BCForth*, while its implementation can be accessed free from the GitHub [16].

A. Core Architecture

Fig. 8 depicts the overall architecture of the Forth language defined in the project *BCForth*. The role and responsibilities of each class in the hierarchy are as follows.

- `TForth` – the base class defining all basic data structures, such as: the data stack represented by `DataStack`, the words' dictionary `WordDict` (`std::unordered_map`), as well as the auxiliary return stack `RetStack`.
  `TForth` defines the `WordEntry`, which is the structure holding all necessary information about a word and kept as a value of each word in the dictionary. `InsertWord_2_Dict` inserts a newly created word to the dictionary, whereas `GetWordEntry` retrieves a word from the dictionary by providing its name as a key; `WordOptional` is returned to cope with situations of non-existing words. Various words are represented as objects from the rich `TWord` family. These have access to the data stack defined in `TForth`. Each word present in the `TForth` dictionary is ready to be executed by calling the `ExecWord` with the `word_name` as its parameter. Hence, `TForth` alone, is sufficient to handle the pre-defined and non-contextual words (i.e. ones that don't need any other tokens from the input stream). This makes `TForth` alone a minimalistic Forth system. `TForth` defines also an auxiliary vector `NodeRepo` to hold objects that need to be present but that do not go to the dictionary of words (Fig. 8). These are e.g. compiled-in literals.
- `TForthInterpreter` – derived from `TForth` is responsible for handling the interpreter mode, in which a stream of `tokens` is processed and executed. Operation of `TForthInterpreter` mostly relies on interpreting the incoming stream of tokens, as integer or floating-point literals (these are distinguished by the dot `.` inside the literal), or as word names to be executed and their optional parameters. However, no new words can be defined (this is a role left for `TForthCompiler`).
- `TForthCompiler` – extends `TForthInterpreter` by providing `the` ability of entering definitions of new words. New words can be entered to the dictionary (Fig. 2) with the defining construction colon-semicolon (`: ;`). For instance,

  ```
  : ACTION DO I . CR LOOP ;
  ```

  defines a new word `ACTION` which upon a call

  ```
  23 0 ACTION
  ```

  prints all values 0-22, each in a new line.
  However, apart from the calls to the words already defined and registered in the dictionary, word definitions can contain nested structural words, such as `IF` … `THEN` ... `ELSE`, `DO` … `LOOP`, etc., as well as the two-stroke `CREATE` … `DOES>` creational pattern, or the `IMMEDIATE` / `POSTPONE` handling modes.
- `TForthReader` – an auxiliary class for converting a text stream, such as a terminal window or a text file, into a stream of Forth's tokens. This is done by text splitting over the white symbols (space, tab, new line), as well as after stripping off the Forth's comments. This way obtained stream of text tokens is fed to the interpreter and/or compiler objects, as described in Section (II.D).

One of the main architectural assumption is a strong separation of the input stream processing components, the token stream processing components, and the word defining objects. In other words, the latter does not bother with any variants of the input and output terminals. On the other hand, the streams of Forth tokens are obtained by the `TForthReader` object. If this is a word definition, tokens are passed to `TForthCompiler`, in order to compile-in a new word. Otherwise, tokens go to `TForthInterpreter` for word(s) execution.

B. Key Data Structures

Details of Forth can be found in many sources [1][4]. Here we focus mostly on the basic data structures and operations which they are used for. Fig. 1 depicts few characteristic operations on the Forth's data stack.

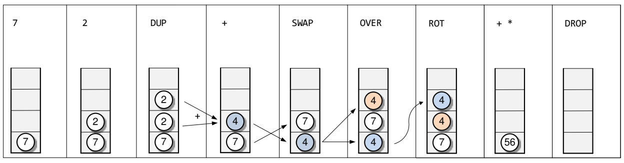


Fig. 1 Examples of the most common stack operations in Forth. All values are entered in the RPN. DUP duplicates the top value of the stack. A binary operator, such as +, removes the two topmost values, performs the operation, and pushes the result onto the stack. SWAP changes order of the two topmost values. OVER copies the second operand and pushes it to the top of the stack. ROT does the rotation of the three topmost stack values. DROP removes the topmost value from the stack.

The operations are straightforward once we recall that all operations are in the Reverse Polish Notation (RPN) [5]. It can be observed that each newly entered value (object) is pushed onto the data stack. Each operation, on the other hand, such as the + operator, or a DUP (duplicate) operation, pops off the necessary number of parameters, performs its specific action, and pushes the result, if there is any (for + this will be the sum, whereas DUP simply duplicates the top value of the stack).

In all of the aforementioned operations an error is thrown if the stack does not contain a number of operands (values) expected by a word. This breaks execution of a word and the special on-error cleaning procedure is launched, after which Forth gets good chances to enter the interpretation mode again, waiting for new commands.

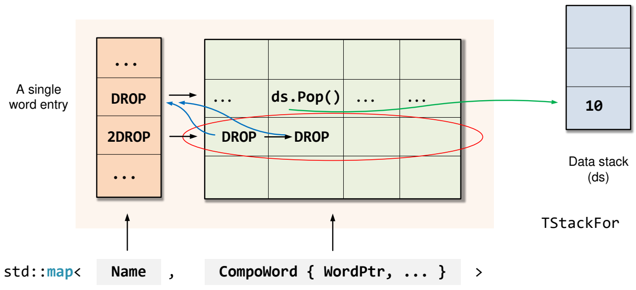


Fig. 2 Forth words are kept in the dictionary data structure, implemented as the C++ `std::unordered_map` with the key being any Name (`std::string`), while definitions are kept in a hierarchical *CompoWord* structures (based on `std::vector` containing pointers to already defined words and other procedures). Frequently, new words call words already present in the dictionary, such as 2DROP which two times calls DROP. Also, the words have an access to the data stack.

The second data structure characteristic to Forth is the word dictionary. Fig. 2 depicts structure of the Forth's dictionary in the *BCForth* implementation that holds definitions of the words, i.e. procedures. Each word is identified by its name.

As shown in Fig. 2, words can access the stack which holds the input and output parameters. Such definition is first scanned by the lexical tokenizer (Fig. 8) to produce valid tokens, such as numeral literals and names of other words. The tokens are then parsed by the Forth compiler and, upon success, new definition is entered to the word dictionary

C. Hierarchy of Words

Fig. 9 depicts hierarchy of word defining classes, which has been already outlined in the general architecture shown in Fig. 8. The roles of the classes in the TWord hierarchy are as follows.

- TWord is a template base class defining functional objects (functors) for the word hierarchy. The F template represents a class that defines all necessary data structures. Currently for this purpose TForth from Fig. 9 is used. Its main functionality, as well as of all its descendant, is *the action* defined by the virtual functional operator (). Naturally, invoking any Forth's word will be translated into calling the corresponding operator (). Hence, the entire TWord hierarchy can be seen as *the command design patter* [8].
- StructuralWord originates the sub-group of the structural words, such as the conditional statement IF … ELSE … THEN, the counted loop DO … LOOP and many more. However, StructuralWord is only a type-holder, whereas the most important function-holder is CompoWord.
- CompoWord defines *the composite design pattern* [8][5] to hold any sequence of Forth's words, also of the same type; such a recursive hierarchy allows composition of nested statements, such as DO … IF … THEN … LOOP, etc.
- IF is an example of *a composite to hold other composites* (similarly other objects in this sub-group). In this case it holds two branches: fTrueBranch representing a set of operation (another composite) chosen if, in the run-time, a condition (a value on the data stack) before the IF statement evaluates to true, and fFalseBranch which stores operations executed on the false condition.
- TValFor and TDataContainer are the two classes to represent a compiled-in value or a container of values, respectively. The type of the stored objects is given by the second template parameter V.
- StackOp – is a variadic template originating the suite of its specializations for defining data stack operations with various number of input and output parameters. For this purpose any function with 0, 1 or 2 input parameters, as well as 0 (void) or 1 return value, can be provided. These are supplied in the form of lambda functions passed to the constructor of the StackOp. Thanks to combination of this variadic template and the lambda functions dozens of stack operations are defined which otherwise required definition of separate classes in the TWord hierarchy [16].
- Dot, Comma, etc. – are examples of specialized system words.

As already mentioned, the key architectural assumption is expression of any Forth's word as the composite pattern, composed of other words, possibly also being composites, and so forth. Such a hierarchical structure provides a flexibility to define language constructions composed of structural statements nested to any depth

D. System Activity

In this section a brief overview of the activity of the Forth's interpreter and compiler are outlined.

The TForthInterpreter class was already outlined in Section (II.A). As shown in Fig. 8, it is directly derived from the base TForth class. Since the main data structures TForthInterpreter inherits from its base, its key role is to execute words from the stream of text tokens, as outlined in the activity diagram shown in Fig. 3.

TForthCompiler is the last and the most complex class in the hierarchy in

Fig. 8. As mentioned, its main responsibility is parsing a word defining stream of tokens, contained in-between the **:** (colon) **;** (semicolon) symbols, and accordingly composing corresponding code of the newly created word.

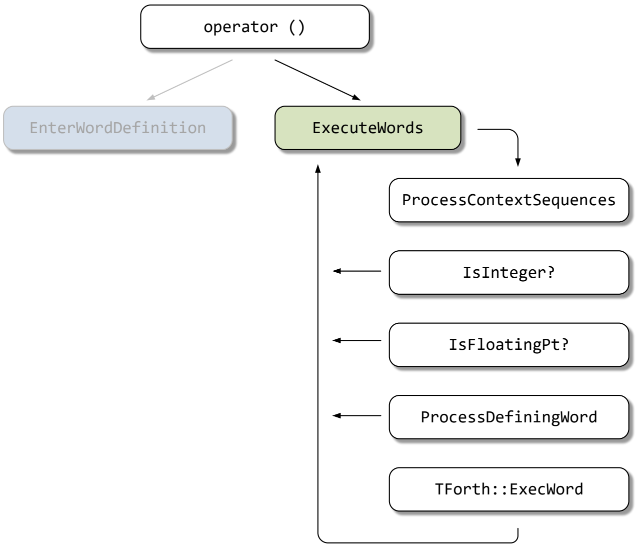


Fig. 3 UML activity diagram of the *TForthInterpreter*. The *ExecuteWords* executes a series of steps after which it is recursively called until the input stream of text tokens is emptied. The *EnterWordDefinition* is the compiler branch

If this operation is successful, the new word is placed in the Forth's dictionary, from which it can be invoked by the interpreter, as well as used in definitions of future words, again processed by the compiler, and so on. Its activity diagram is shown in Fig. 5.

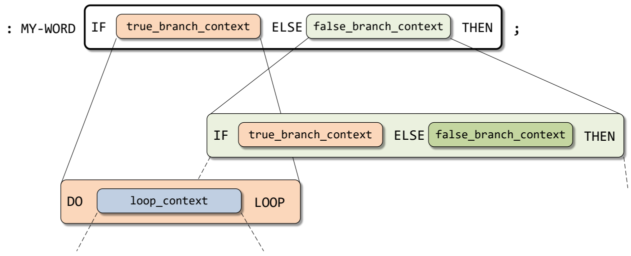


Fig. 4 Changing context concept. Each word, as well as each branch of a structural construction such as `IF` … `ELSE` … `THEN`, `DO` … `LOOP`, etc. has its own context implemented with its own composite *CompoWord*. Each such object has links to other words, also other *CompoWords*, and so on. The entire structure is parse by a successive recursive call to the parsing procedure

Fig. 4 depicts an example of the nested structure constructions. The key observation is that each sub-branch opens *a new context*, which can be treated as a separate sub-word construction, and so on. This creates a hierarchical composition which can be processed in a recursive manner – at each level the sub-branch is processed *independently* as a separate sub-word and in *its own context*.

Finally, the remaining Forth words are defined in separate Forth modules. These are special classes (Fig. 9) to enter word definitions for various domains, such as floating-point, string & memory processing, and from different sources, such as hard coded, string or file stream. The pure abstract root `TForthModule` starts their class hierarchy

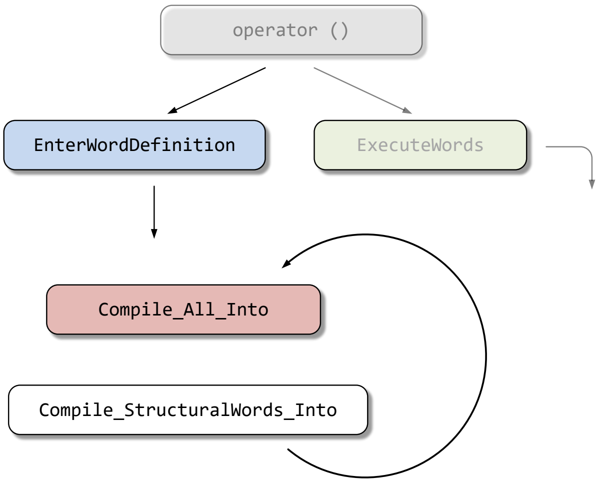


Fig. 5 Activity diagram of *TForthCompiler*. *EnterWordDefinition* implements the principal functionality of the Forth compiler – parsing word defining stream of text tokens and constructing the corresponding implementation. To process structural statements, which can be nested to any depth, each structural statement enters into a new context represented by a separate composite object. Processing is done by recursive calls of the *Compile_All_Into* function until the entire defining stream is processed. *Compile_StructuralWords_Into* processes the structural statements such as conditional IF, DO, etc. in interaction with the structural words of the *TWord* hierarchy

III. FORTH ENDOWED WITH THE COROUTINES

Existence and roles of functions, or routines, in computer programs are ubiquitous and well known. However, there is a special type of a routine called a coroutine, which can suspend its execution preserving its state to be resumed later [9], as shown in Fig. 6.

For such functionality coroutines need to have associated memory to store local data and the resumption point. In this respect there are two groups: stackfull and stackless coroutines. Modern C++20 provides the framework and mechanisms for the latter [7][3]. That is, they suspend execution by returning to the caller and the data that is required to resume execution is stored separately from the stack. This allows for sequential code that executes asynchronously e.g. to handle non-blocking I/O without explicit callbacks, allows for the so called lazy-computations e.g. to generate infinite series of values, but most of all it allows for cooperative multitasking purely on the Forth platform. The latter is very useful feature especially on small and resource constraints platforms that nevertheless require the kind of multitasking [2]. Forth built in coroutines allow for such operation in much more lightweight way compared to the preemptive multitasking. Hence, coroutines are a unique feature of *BCForth*.

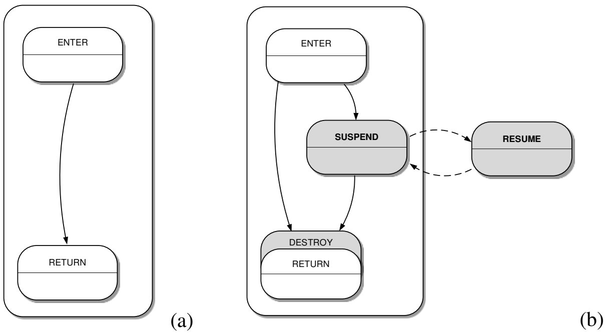


Fig. 6 State diagram of an ordinary routine (a) and a coroutine (b). The latter can also suspended, preserving its state, to be resumed later. This allows for async operations or lightweight threading

The main proposed idea is to introduce new Forth words, which will operate as the stackless co-routines (however, they have an access to the Forth's stack). For this purpose a new word named **`CORO`** is proposed, which if put after a word's definition, makes it a coroutine (this is similar to the `IMMEDIATE` post word). In the Forth's nomenclature we propose to call them *co-words*. For instance, the following defines the word `FIBER_0` that does `XOR` of the first cell in a buffer `BUF`, then reads the second cell from that buffer and pushes it onto the Forth's stack

```
: FIBER_0 BUF @ 0xAB XOR BUF !   0x02 BUF + @
                ;   CORO  [155]
```

However, `CORO` with the optional parameter `[155]` makes it a Forth's coroutine that toggles some bits, and suspends after 155 ms, or terminates if the second cell in `BUF` is not 0. This is possible thanks to the `CORO_Frame` and `FiberTask<T>` C++ coroutine structure that operates as a wrapper around any `WorkerWord`, such as `FIBER_0`, in our example, while `time_slice` becomes 155. An outline of `CORO_Frame` looks as shown in Algorithm 1. `FiberTask<T>` in line [1] is a structure with the nested class `promise_type`, as required by the C++20 framework [7]. On the other hand, `GetTimePoint` in lines [3,5,10] does the time management, resulting with the suspend via `co_await` in [9].

The next proposed new word is **`COYLD`** (from *co-yield*) that suspends a given word leaving its value on the top of the Forth's stack. Thanks to this, the value generating words can be defined. With its help the `CO_RANGE` word has been created which, upon each call, generates and pushes onto the Forth's stack consecutive values from a predefined range. For example `10 20 2 CO_RANGE` creates a generator of values 10 to 20 with step of 2. Then each call to `CO_RANGE` leaves 12, 14, …, 18 on the stack.

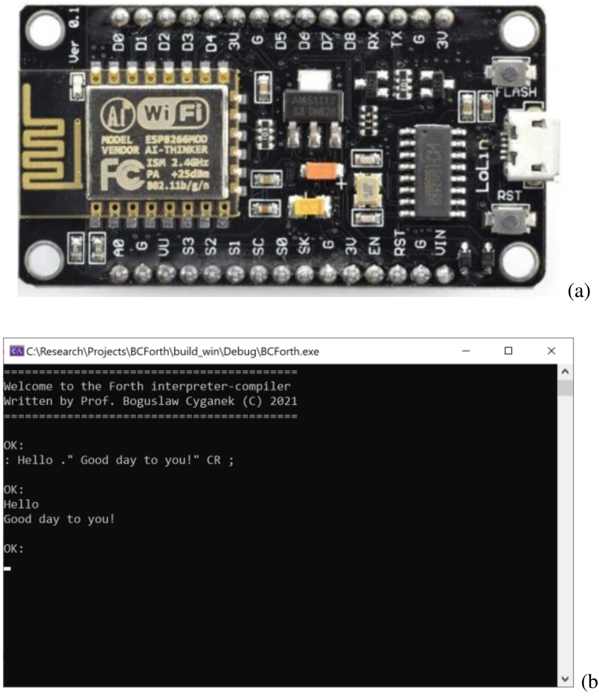
The last from the proposed words is the **`COSUS`** (from *co-suspend*). It suspends a Forth's word at its point of call, from which that word will resume if called again (naturally, an 'ordinary' Forth word would start from the beginning).

IV. System Deployment and Experiments

The complete C++ implementation of *BCForth* with exemplary Forth programs is available from the GitHub [16]. This is *a multi-platform header only library* aimed at Linux/Unix, Windows, and MacOS. It was successfully built and deployed on the following platforms:

1. PC computer with Linux Ubuntu 18.4 and 20.4, run on laptop Dell Precision 7710. Compiled with the gcc version 10 and 11. The latter allows co-routines.
2. PC computer with Windows 10 run on laptop Dell Precision 7760. Compiled with the Microsoft Visual C++ 2019 v. 16.9.2, as well as MV 2022 v. 17.2.6.
3. Embedded system NodeMCU v3 with the 32-bit RISC ESP8266 microprocessor [17][18], controlled by the 80 MHz clock (based on Tensilica Diamond Standard 106Micro architecture). The system equipped with the 32 KB instruction memory and 80 KB data RAM. The system contains built-in Wi-Fi, 10 GPIO ports, ADC converter and USB-UART CH340 link, allowing also external programming. Built in the PlatformIO Arduino equipped with the gcc version 10. This is an example of a IoT tiny platform with its own system but yet without co-routines.

Fig. 7(a) depicts the NodeMCU board, while *BCForth* run in the interactive mode in the terminal window is shown in Fig. 7(b).



```
==========================================
Welcome to the Forth interpreter-compiler
Written by Prof. Boguslaw Cyganek (C) 2021
==========================================

OK:
: Hello ." Good day to you!" CR ;

OK:
Hello
Good day to you!

OK:
```

Fig. 7 Embedded system NodeMCU v3 with the 32-bit RISC ESP8266 microprocessor 32KB+80KB RAM, 80 MHz clock (a). *BCForth* running in the terminal window (b)

Although both Linux and Windows 10 allowed for a complete implementation, special attention deserves the third platform which is a tiny NodeMCU v3 embedded systems with only the 32 KB instruction memory and 80 KB data RAM.

Nevertheless, with some minor modifications, it was also possible to run *BCForth*. This shows that despite C++ the footprint of the *BCForth* can be as small as to fit to the small (and cheap) embedded platforms and/or IoT systems.

However, even more important is fast time (approx. three weeks) of *BCForth* system tuning to the new platform done by Mr. W. Gałecki & Ms. K. Rapacz, students of the 1st year of the graduate studies Electronics & Telecommunication, as a completion of their project to the Systems Design and Modeling Methodologies classes under author's supervision at the AGH University of Science and Technology. This proves that *BCForth* implementation is straightforward for all persons with at least medium competitions in the modern C++ programming, as well as that it can be easily deployed on similar tiny embedded frameworks. This also adds the teaching aspect of the presented system and, hopefully, can be used with educational and technical benefits by a broader group of students and enthusiasts of embedded systems

V. Conclusions

In this paper a novel and free Forth language platform *BCForth* [16], aimed at embedded systems of various sizes, is proposed. The main advantage of Forth is coexistence of the compiler and interpreter that allows for direct communication with a user and easy composition of new words (procedures). Unique *BCForth* features are as follows: (i) modular C++20 based implementation, (ii) implementation of coroutines for *async* operations and lightweight multithreading, (iii) educational/teaching platform for students of electrical engineering faculties. Envisioned things to do are: (i) modules with new words (e.g. file operations, graphics, etc.), (ii) GUI for Forth development and debugging, (iii) auto setup for easier deployment on the limited footprint platforms. We are deeply convinced that this novel implementation of Forth will be beneficial for embedded systems, as well as in education and further popularization of Forth and C++.

References

[1] Brodie L.: Thinking Forth. A Language and Philosophy for Solving Problems, Creative Commons, 2004.

[2] Belson B., W. Xiang, J. Holdsworth and B. Philippa, C++20 Coroutines on Microcontrollers – What We Learned, IEEE Embedded Systems Letters, vol. 13, no. 1, pp. 9-12, 2021.

[3] Belson B., et al.. A Survey of Asynchronous Programming Using Coroutines in the Internet of Things and Embedded Systems. ACM Trans. Embed. Comput. Syst. 18/3, 2019.

[4] Conklin E.K., Rather E. D.: Forth Programmer's Handbook, FORTH Inc. 2010.

[5] Cyganek B.: Introduction to Programming with C++ for Engineers. Wiley-IEEE Press, 2021.

[6] Dunkels A., Schmidt O., Voigt T., Muneeb A. Protothreads: simplifying event-driven programming of memory-constrained embedded systems. 4th international conference on Embedded networked sensor systems (SenSys '06). ACM, 29–42, 2006.

[7] https://en.cppreference.com/w/cpp/language/coroutines

[8] Gamma, E.; Helm, R.; Johnson, R.; Vlissides, J.: Design Patterns: Elements of Reusable Object-Oriented Software, Addison-Wesley Professional, 1994.

[9] Knuth D. E. The art of computer programming, Vol. 1: Fundamental algorithms (3rd. ed.), Addison-Wesley, 1997.

[10] Pelc S.: Programming Forth. MicroProcessor Engineering Limited, 2005.

[11] Rather E. D., Colburn Donald R., and Moore Charles H.: The evolution of Forth. History of programming languages-II. Assoc. for Comp. Machinery, New York, USA, 625–670, 1996.

[12] Rather E. D.: Forth Application Techniques, 6th edition, FORTH Inc. 2019.

[13] https://en.wikipedia.org/wiki/Charles_H._Moore#cite_note-2

[14] https://forth-standard.org/

[15] https://theforth.net/

[16] https://github.com/BogCyg/BCForth

[17] https://en.wikipedia.org/wiki/ESP8266

[18] https://en.wikipedia.org/wiki/NodeMCU

[19] https://en.wikipedia.org/wiki/Forth_(programming_language)

[20] https://gforth.org/

[21] https://www.forth.com/

[22] http://git.annexia.org/?p=jonesforth.git;a=summary

[23] https://awesomeopensource.com/projects/forth

[24] https://cppreference.com

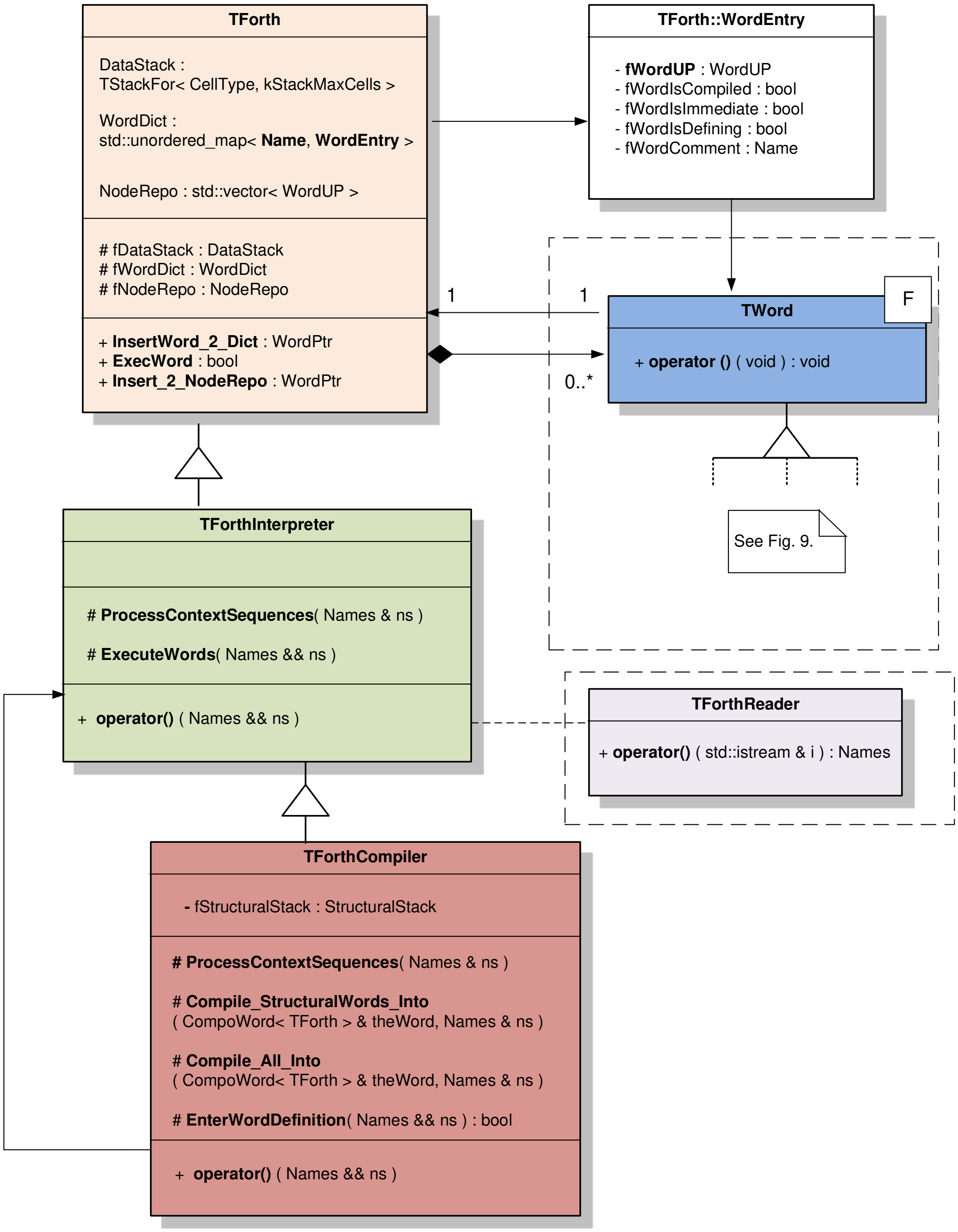

TForth
DataStack :
TStackFor< CellType, kStackMaxCells >
WordDict :
std::unordered_map< **Name**, **WordEntry** >
NodeRepo : std::vector< WordUP >
# fDataStack : DataStack
# fWordDict : WordDict
# fNodeRepo : NodeRepo
+ **InsertWord_2_Dict** : WordPtr
+ **ExecWord** : bool
+ **Insert_2_NodeRepo** : WordPtr
TForth::WordEntry
- **fWordUP** : WordUP
- fWordIsCompiled : bool
- fWordIsImmediate : bool
- fWordIsDefining : bool
- fWordComment : Name
1
1
0..*
TWord
F
+ **operator ()** (void) : void
See Fig. 9.
TForthInterpreter
# **ProcessContextSequences**(Names & ns)
# **ExecuteWords**(Names && ns)
+ **operator()** (Names && ns)
TForthReader
+ **operator()** (std::istream & i) : Names
TForthCompiler
- fStructuralStack : StructuralStack
# **ProcessContextSequences**(Names & ns)
# **Compile_StructuralWords_Into**
(CompoWord< TForth > & theWord, Names & ns)
# **Compile_All_Into**
(CompoWord< TForth > & theWord, Names & ns)
# **EnterWordDefinition**(Names && ns) : bool
+ **operator()** (Names && ns)

Fig. 8 Architecture of the *BCForth* system. The main branch is composed of three classes: *TForth*, *TForthInterpreter* and *TForthCompiler*. These use the hierarchy of word nodes, originated from the *TWord* base (*Names* denotes a collection of text tokens). The input/output operations are interfaced by the *TForthReader* class, which transforms an input stream (terminal or a file) to a series of tokens.

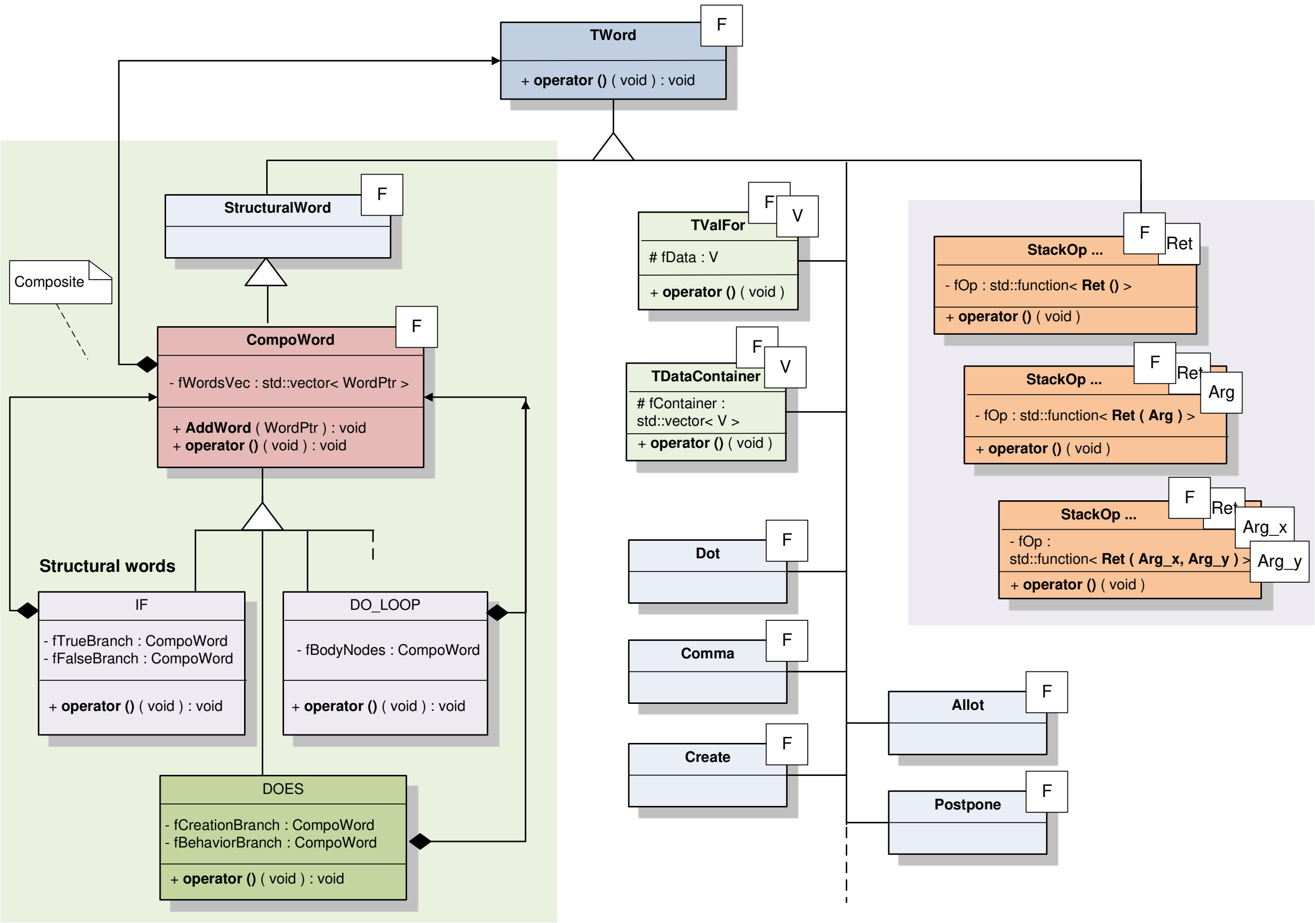

TWord
F
+ operator () (void) : void
StructuralWord
F
Composite
CompoWord
F
- fWordsVec : std::vector< WordPtr >
+ AddWord (WordPtr) : void
+ operator () (void) : void
Structural words
IF
- fTrueBranch : CompoWord
- fFalseBranch : CompoWord
+ operator () (void) : void
DO_LOOP
- fBodyNodes : CompoWord
+ operator () (void) : void
DOES
- fCreationBranch : CompoWord
- fBehaviorBranch : CompoWord
+ operator () (void) : void
TValFor
F
V
# fData : V
+ operator () (void)
TDataContainer
F
V
# fContainer :
std::vector< V >
+ operator () (void)
Dot
F
Comma
F
Create
F
Allot
F
Postpone
F
StackOp ...
F
Ret
- fOp : std::function< Ret () >
+ operator () (void)
StackOp ...
F
Ret
Arg
- fOp : std::function< Ret (Arg) >
+ operator () (void)
StackOp ...
F
Ret
Arg_x
Arg_y
- fOp :
std::function< Ret (Arg_x, Arg_y) >
+ operator () (void)

Fig. 9 Hierarchy of classes defining the Forth words. The composite design pattern, implemented with the *CompoWord* branch, constitutes the main building block of all words registered to the Forth dictionary. The next branch constitute system specific words, such as *Dot* or *Create*. The *StackOp* branch, implemented as a variadic template and its specializations, is responsible for majority of the data stack operations, such as arithmetic and logical operations.

Algorithm 1. Scheme of the *CORO_Frame* routine.

```
template < typename T, auto time_slice, auto WorkerWord >
FiberTask< T > CORO_Frame ( auto worker_load ) {
    auto tp0 = GetTimePoint(); // coroutines suspend on time elapsed
    for( ;; ) {
        // embed/call worker word WorkerWord with worker_load
        // if done, then co_await or break for co_return

        if( GetTimePoint() - tp0 > time_slice ) {
            co_await std::suspend_always{}; // suspend if time elapsed
            tp0 = GetTimePoint();
        }
    }
    co_return -1;
}
```

Joint 42nd IEEE Software Engineering Workshop and 9th International Workshop on Cyber-Physical Systems

THE IEEE Software Engineering Workshop (SEW) is the oldest Software Engineering event in the world, dating back to 1969. The workshop was originally run as the NASA Software Engineering Workshop and focused on software engineering issues relevant to NASA and the space industry. After the 25th edition, it became the NASA/IEEE Software Engineering Workshop and expanded its remit to address many more areas of software engineering with emphasis on practical issues, industrial experience and case studies in addition to traditional technical papers. Since its 31st edition, it has been sponsored by IEEE and has continued to broaden its areas of interest.

One such extremely hot new area are Cyber-physical Systems (CPS), which encompass the investigation of approaches related to the development and use of modern software systems interfacing with real world and controlling their surroundings. CPS are physical and engineering systems closely integrated with their typically networked environment. Modern airplanes, automobiles, or medical devices are practically networks of computers. Sensors, robots, and intelligent devices are abundant. Human life depends on them. CPS systems transform how people interact with the physical world just like the Internet transformed how people interact with one another.

The joint workshop aims to bring together all those researchers with an interest in software engineering, both with CPS and broader focus. Traditionally, these workshops attract industrial and government practitioners and academics pursuing the advancement of software engineering principles, techniques and practices. This joint edition will also provide a forum for reporting on past experiences, for describing new and emerging results and approaches, and for exchanging ideas on best practice and future directions.

TOPICS

The workshop aims to bring together all those with an interest in software engineering. Traditionally, the workshop attracts industrial and government practitioners and academics pursuing the advancement of software engineering principles, techniques and practice. The workshop provides a forum for reporting on past experiences, for describing new and emerging results and approaches, and for exchanging ideas on best practice and future directions.

Topics of interest include, but are not limited to:

- Experiments and experience reports
- Software quality assurance and metrics
- Formal methods and formal approaches to software development
- Software engineering processes and process improvement
- Agile and lean methods
- Requirements engineering
- Software architectures
- Design methodologies
- Validation and verification
- Software maintenance, reuse, and legacy systems
- Agent-based software systems
- Self-managing systems
- New approaches to software engineering (e.g., search based software engineering)
- Software engineering issues in cyber-physical systems
- Real-time software engineering
- Safety assurance & certification
- Software security
- Embedded control systems and networks
- Software aspects of the Internet of Things
- Software engineering education, laboratories and pedagogy
- Software engineering for social media

TECHNICAL SESSION CHAIRS

- **Bowen, Jonathan,** Museophile Ltd., United Kingdom
- **Hinchey, Mike**(Lead Chair), Lero-the Irish Software Engineering Research Centre, Ireland
- **Szmuc, Tomasz,** AGH University of Science and Technology, Poland
- **Zalewski, Janusz,** Florida Gulf Coast University, United States

PROGRAM COMMITTEE

- **Ait Ameur, Yamine,** Toulouse Institute for Research in Computer Science, France
- **Banach, Richard,** University of Manchester, United Kingdom
- **Challenger, Moharram,** University of Antwerp, Belgium
- **Cicirelli, Franco,** DIMES Università della Calabria, Italy
- **Gomes, Luis,** Universidade NOVA de Lisboa, Portugal
- **Gracanin, Denis,** Virginia Tech, USA

- **Havelund, Klaus,** Jet Propulsion Laboratory, USA
- **Hsiao, Michael,** Virginia Tech, USA
- **Karaduman, Burak,** University of Antwerp, Belgium
- **Katwijk, Jan van,** TU Delft, Netherlands
- **Pullum, Laura,** Oak Ridge National Laboratory, USA
- **Sekerinski, Emil,** McMaster University, Canada
- **Sojka, Michal,** Czech Technical University in Prague, Czech Republic
- **Trybus, Leszek,** Rzeszow University of Technology, Poland
- **Vardanega, Tullio,** University of Padua, Italy
- **Velev, Miroslav,** Aries Design Automation, USA

DOI: 10.15439/2022F287
ISSN 2300-5963 ACSIS, Vol. 31

Integrated Checklist for Architecture Design of Critical Software Systems

Adela Bierska, Barbora Buhnova and Hind Bangui
Faculty of Informatics, Masaryk University
Brno, Czech Republic
{bierska, buhnova, hind.bangui}@mail.muni.cz

***Abstract*—With the advancement of digitalization, critical information infrastructures, such as intelligent energy distribution, transportation, or healthcare, have opened themselves towards intelligent technological opportunities, including automation of previously manual decision making. As a side effect, the digitalization of these infrastructures gives rise to new challenges, especially linked to the complexity of architecture design of these infrastructures, to later support necessary software quality and safeguard the systems against attacks and other harm. To support software architects in the design of these critical software systems, well structure architectural knowledge would be of great help to prevent the architects from missing some of the crucial concerns that need to be reflected with built-in architectural mechanisms, early during architecture design.**

Given the narrow scope of existing guidelines, with the need of browsing and combining multiple sources, this paper proposes an integrated checklist to cover the breath of architectural concerns for the design of critical software systems, covering the need for built-in mechanisms to prevent, detect, stop, recover from and analyse intentional as well as unintentional threats to system dependability. Contrary to existing guidelines that typically focus on runtime incident handling, our checklist is to be used during architecture design to ensure that the system has built-in mechanisms to either handle the incidents automatically or include the right mechanisms to support the runtime incident handling.

***Index Terms*—Software architecture, design checklist, critical information infrastructure, dependability**

I. Introduction

CRITICAL information infrastructures could be understood as digital and vital systems that require immediate attention and protection in modern cities (e.g., intelligent transportation) because they contribute in the improvement of the quality of life and sustainable development of our society. Over the past decades, critical infrastructures in various domains of human life have become largely digitized, stressing achievement of system security, resilience, reliability and other characteristics of failure-free and dependable operation [1], [2]. However, although these systems have become highly software intensive, software architecture experts have often not been involved in the design of these systems, which is why the operators of these systems are seeking software architecture expertise ex-post to evaluate and improve software architecture design of these systems.

While software architects have access to numerous standards and guidelines for system design and auditing in concrete domains of critical infrastructures, e.g., CIPSEC [3] or Cybersecurity Certification [4], there is no general overview of design guidelines they shall consider, which is leaving them with substantial risk that they might miss some crucial consideration. More so that existing standards and guidelines disproportionately more focus on hardware considerations, which might make it even more likely for software architect that they miss some software-architecture related aspects.

To address this gap, the aim of this paper is to propose the creation of an integrated checklist supporting software architects in the design of critical software systems. The checklist is meant to cover guidelines that help the architect to design buit-in mechanisms to improve the dependability of the designed system. Given the existence of various low-level tactics and patterns, e.g. for availability, reliability or security in specific types of systems [5], [6], the suggested critical infrastructure design guidance shall be high-level and integrative, emphasizing the specifics of critical information infrastructures that might otherwise be missed. This paper introduces the reader to the context of critical infrastructures and software incidents and explains the checklist creation process and its usage, with additional supplementary material available for download at [7].

The structure of the paper is as follows. After the introduction, the context of the topic together with the state of the art and related work is presented in Section II. Section III lays down the design considerations guiding the design of the checklist. Section IV details the methodology used to create the checklist, after which the resulting checklist is presented in Section V. Additionally, supplementary material is available at [7], containing the full guidelines classification data, detailed guidelines descriptions, and a demonstration of the checklist in a real-life context.

II. Critical Systems and Infrastructures

There are numerous definitions of the term *critical (information) infrastructure (CI)*[1] in the literature from legal, political, technical, economical, geographical or social perspectives [8]. The German Federal Ministry of the Interior,

This research was supported by ERDF "CyberSecurity, CyberCrime and Critical Information Infrastructures Center of Excellence" (No. CZ.02.1.01/0.0/0.0/16_019/0000822).

[1] The word *information* within *critical information infrastructures* is being used to emphasize reference to ICT enhanced critical infrastructures.

for instance, defines critical infrastructures as: "organizational and physical structures and facilities of such vital importance to a nation's society and economy that their failure or degradation would result in sustained supply shortages, significant disruption of public safety and security, or other dramatic consequences" [9]. Overall, there is an agreement that critical infrastructure is an infrastructure that is needed to keep other major technical and/or social systems running or which is needed to provide goods or services that are considered vital to the functioning of modern society [8].

1) Challenges in Architecture Design of CIs: Quality engineering is an inherent goal of software architecture design in any domain, driving the main software architecture activities, such as encapsulation, partitioning, or legacy components integration, using techniques such as isolation, redundancy allocation, or distribution [5], [10].

While in more wide-spread and popular domains, such as enterprise systems and web applications, many architectural patterns and styles exist [11], [12], integrated guidance for critical infrastructures is so far missing.

Obviously, various low level recommendations exist for specific infrastructures and quality attributes to be optimized [5], [6], where the infrastructures differ according to the local context of each nation, and the quality attributes differ according to the infrastructure, with many specific attributes not covered in existing software architecture literature (e.g. absorbtion as the ability to absorb the effects of system failures and thus minimize the consequences, or preparedness as the ability to withstand the expected crisis situations). This guidance, besides being isolated and localised, is moreover scarce and often provided in terms of *answers* rather than *questions to be asked*, which makes the critical infrastructure design process highly error-prone.

2) SW Architecture Design Checklist: Simple checklists can help reduce human error dramatically. As emphasized by Ivar Jacobson when advocating for software project checklists [13], "Neil Armstrong had a checklist printed on the back of his glove to ensure he remembered the important things as he made history as the first person to walk on the moon, so why not utilize them to keep software projects on track." Checklists in software engineering are not a new concept, being employed to facilitate software development processes by efficiently guiding industry experts and professional developers.

III. Design Considerations for Dependable CIs

Critical infrastructures are by definition safety-critical and often handle sensitive, secret, or in other way valuable data [14]. Consequently, they are an attractive target for attackers and need more robust protection.

Software-intensive critical infrastructures are highly complex as they operate with various technologies and have to be highly reliable. While designing the system, an architect may forget numerous basic features and cause vulnerabilities. This could endanger people or cost money. Our checklist is designed for these infrastructures to help its architects to meet the crucial safety and security standards and meet their high reliability expectations.

In this section, we set the foundations for the checklist, discussing its scope and analysing the types of incidents it shall cover, supporting the software architecture mechanisms to prevent, detect and handle them.

A. Domains

The most common CIs around the world, according to CIPSEC [15] are in the domains of health, energy, transportation, finance, food, water, and civil administration. Each of these CIs faces a different set of threats and safety concerns, but in their essence, the infrastructures need to meet very high standards that are very similar among them. In the presented version of the checklist, we thus focus on the design concerns that need to be reflected by all of them.

CI domains are usually interdependent [16]. For example, all of those mentioned above depend on the energy sector as they use software systems and machines for their operations. The energy sector relies on water supply used for cooling, which is impossible without transporting material within the infrastructure [15]. This can cause a domino effect throughout the whole system [17]. That is why it is necessary to stop an error as soon as possible and ensure the functioning of the most critical parts.

B. Types of Incidents

A software incident is an event that brings the system to an unwanted, anomalous state [18]. Incidents in CIs can be catastrophic and endanger human lives or the economy. Therefore software should be prepared to prevent and stop them.

We can classify incident causes in different ways. They can be intentional or unintentional, man-made or caused by natural disasters, caused by an error in code or hardware [19].

One of the most significant challenges of software dependability is, for the time being, unknown threats and attacks. New types of vulnerabilities are discovered every day [20], and natural disasters are often also unpredictable. Due to this, the software architect may not be aware of perils that will be ordinary for the system in a few years, months, or even days. Therefore we should prepare the system universally and not rely on a concrete list of possible threats.

1) Natural-Disaster Causes: Natural incidents are caused by natural disasters like floods, tornados, earthquakes, etc. [21] They can cause damage to the system hardware and therefore impact the correct functioning of the software. From the software point of view, we can also include in this category the incidents triggered by hardware errors (i. e., outage of parts) or by other damage to the hardware (i. e., incompetent manipulation or militarized attack). Natural disasters can be unpredictable, devastating, and cause a domino effect on other domains of CIs [21]. This enormously complicates the possibilities of testing the system's behavior during these incidents [22].

2) Man-Made Causes: By man-made incidents, we mean inadvertent mistakes made by a user or a premeditated attack from an attacker. In contrast to natural incidents, we can prevent many of these.

a) Accidental Errors: There are many reasons why users can make mistakes while using the system. For example, they may not know how to use the system, they can be tired and careless, or the system may be confusing [23]. In any case, this inappropriate usage should not endanger the correct functioning of the system.

b) Deliberate Attacks: Deliberate attacks on CIs occur increasingly often [24]. Main objectives of attackers are [25]:

- **Corruption of information:** Attackers try to change or damage data stored in the system or corrupt communication.
- **Denial of service:** Attackers try to overload or disrupt the system to become unavailable for authorized users.
- **Disclosure of information:** Attackers try to obtain private data or publish them to unauthorized entities.
- **Theft of resources:** Attackers try to access and misuse the system resources or provide them to unauthorized entities.
- **Physical destruction:** Attackers try to cause physical damage using the system.

The software should be prepared for all kinds of attacks, prevent them, and ensure safety in case of malicious usage. In contrast to accidental errors, deliberate attacks may last longer and be more complex. Attackers systematically conceal their activity, so detecting such incidents can be tricky.

C. The Role of Checklists

A checklist is a structured list of requirements or steps needed to achieve the given goal. It can have various structures [26] but should be brief and synoptical to minimalize the cognitive load of its usage [27]. When designing a software system, we can come across checklists, for example, in chosen standards. Checklists usually contain a list of conditionals that need to be fulfilled to meet the given standard [26].

In this work, we were inspired by checklists used by experts to facilitate their decision-making, e.g., in healthcare. Here, it can help us on two different levels – to make our decision (or diagnosis) or to check the correctness of our already-made decision [27]. Experts often tend to evaluate the system based on experienced patterns, which, however, may be superficial. In this case, the checklist reminds them of essential points that they should consider [28].

IV. Methodology

To ensure the comprehensiveness and completeness of the checklist, we had to collect guideline sources to base the checklist on. First, we gathered 32 standards related to software or software-intensive CIs based on criteria in Section IV-A, filtered them to select the representatives with complete coverage of the others, prioritizing freely available sources so that the architect can be pointed to them from the checklist [29], [30], [31], [32], [33], [34], [35], [36]. We collected all requirements meeting our inclusion and exclusion criteria (see section IV-B) from these standards and extracted the most common categories (i.e., Authorization, Authentication, Data protection, Logging, Input and Output, Network, Safety Ensuring, Backups, Encryption, and Third-Party Components).

We added categories to cover incident phases described in Section IV-C, i.e., Access Control, Anomaly Detection, Phenomenon Evaluation, Stopping from Propagating, Self-Adaptiveness, and Evidence. For each category, we went through existing studies and other sources [37], [38], [39], [40], [41], [42], [43], [44], [45], [46], [47], [48], [49], [50], [51], [52], [53], [54], [55], [56], [57] to examine the completeness of the coverage by the standards and complement the missing pieces. All these recommendations were again validated against our inclusion and exclusion criteria.

Finally, we revised categories, some of them were merged or removed, but we also added some to facilitate orientation within the checklist.

A. Sources of Guidelines

The checklist is based on two categories of sources: standards and other recommendations. Each has slightly different conditions for inclusion, but both must be primarily software-related.

Standards must be official and published by an approved organization, government department or peer-reviewed publisher. They should contain a set of rules and aims which ensure software security or safety. They can be designed for critical infrastructure in general, but they have to include a part aimed concretely at software.

When choosing standards, we primarily focused on their domain subsumption to critical infrastructures. Some of them are universal to all software. In that case, they should mention critical systems in their scope or focus on technology commonly used in our target domains.

By recommendations, we mean the research studies and books that software architect can study and abide by to accomplish system security and reliability. It is out of the scope of the checklist to go in a fine detail, it shall rather focus on a general guidance with great coverage in terms of the breath, not depth.

B. Scope

To fully utilize the advantages of the checklist form (e.g., briefness, coverage), we have to set strict inclusion and exclusion criteria of guidelines. The checklist should be compact but cover all identified categories.

1) Inclusion Criteria: Every guideline should be generally usable within CI systems. There may be exceptions for particular systems, but in general, all guidelines should focus on the design of dependable (safe, reliable and secure) critical systems. All guidelines must be relevant to the software architect and propose design improvement to software systems. The included guidelines shall reflect that not only do we try to prevent problems or incidents, but we are also designing systems to be able to operate safely in presence of such problems and incidents, i.e. we have to prepare the system to withstand accidents that have penetrated. Even with the robust, firmly secure architecture, we cannot presume that incidents are impossible [58].

2) Exclusion Criteria: (a) Human Resources: Many cyber-attack related studies support our incident handling phases division but aim more at management and incident planning [59], [60], [61]. Human resources and management are also excluded from the checklist scope. *(b) Hardware:* Considering that our guidelines focus on software, we exclude all purely-hardware related items. Hardware solutions are often irrelevant for software designers and strongly depend on their domain and chosen technologies. *(c) Coding Practices:* The choice of the programming language [62], frameworks and libraries can significantly affect the complexity of the development process. However, that shall is the task of the software architect to set constraints on such low level of detail. Therefore purely coding practices are excluded from the scope.

C. Incident-Handling Phases

The process of incident handling, which needs to be supported by the architecture design checklist, can be structured in multiple phases. In this work, we use the phases inspired by the *Computer Security Incident Handling Guide* [18], which sets them for handling already running incidents. In our case instead, we employ the phases to structure the guidelines *to design software architectures towards preparedness for these phases* making systems more robust and dependable (reliable and secure). The phases are:

- **Prevention:** The prevention of incidents strongly depends on the overall environment of the system. To specify guidelines for this phase, we have to consider all potential sources of failure and use appropriate architectural guidelines to safeguard all the possible entry points for the incidents to enter the system, taking both external and insider attacks, as well as events such as natural disasters into consideration.
- **Detection:** The primary aim of this phase is to design the system with built-in mechanisms to detect a running incident. That is to detect anomalies in the behaviour and discover intruders who might steal data without changing system behaviour. With the correct Detection in place, the system can switch on time to the Containment phase. The secondary (not less important) purpose of the Detection phase is to classify the fault and find its source. These are essential for stopping the fault from propagating and its following elimination.
- **Containment:** This phase covers tactics prepared for an immediate reaction to the detected incident, mainly stopping the problem from propagating and ensuring safety. While sometimes we need to stop the fault as soon as possible, in other cases we may only consider slowing down the attacker or using sandboxing. In any case, it is essential to bypass critical parts of software, especially those responsible for ensuring safety. Furthermore, the checklist shall motivate the architects to identify the must-work functionalities, on which the safety depends, and ensuring these parts work.
- **Recovery:** Right after getting the fault under control, the system shall have the right mechanisms to start the recovery process. It should be able to isolate and remove all infected and suspicious parts, as damaged system is more vulnerable to recurrent failures and attacks. Pre-incident preparation is fundamental as we need to compare states before and after and restore backup data.
- **Post-Incident Analysis:** An essential part of the post-incident analysis is forensic investigation, so the system should be designed and prepared for it before the incident occurs. Pieces of evidence that the system stores should be admissible at court of law and comply with local constraints to data monitoring and storing (e.g., GDPR).

The phases are not strictly separated, some guidelines can support multiple phases. For example, prevention must remain stable during the incident to impede collateral attack. Collection of evidence starts with recognizing the incident [63]. This classification of phases ensures its clear coverage of the security and reliability of the designed system. It also provides better possibilities for synoptical structure.

V. Integrated Checklist for Designing CIs

The primary purpose of checklists, in general, is to keep track of particular processes or units. They remind us of all steps we have to fulfill to complete the task and record every subtask we have already accomplished. If well designed, we can use them for self-evaluating without a checkup from another entity.

Our checklist provides this feature of self-evaluation of software systems by listing the must-haves for handling particular phases of security incidents and offering relevant standards and sources. After going through all of the provided guidelines, the software architect should be able to make sure that the architecture contains buit-in mechanisms to reflect all the given concerns.

All systems are unique and struggle with various issues. The checklist should cover the plentifullest amount of them and stay adaptive and concrete. Therefore it offers the available guidelines fulfilling our scope, in its breath, and also allows the architect to choose which are relevant.

Last but not least profitable feature of the checklist is facilitating communication in the development team. It should be readable by all team members independently on whether they are creators.

A. Structure

The structure has to observe typical qualities of checklists: briefness and clarity. Its main aim is to classify, sort external sources, and show only short descriptions and references. Any user should be able to browse all offered guidelines without previous knowledge of the checklist.

The form of the checklist is variable. Initially, the designer gets all the sources categorized by incident phases and tags.

TABLE I
PHASE 1: PREVENTION

GUIDELINE	TAGS	SOURCES
Data Protection		
Saved data are secured.	Data Protection, Encryption, Network	[32], [37], [38]
Network communication (internal too) is secured.	Data Protection, Network, Authentication	[32], [29], [33], [37], [36], [38]
Authorization		
Each entity has a specific role within the system with minimal necessary rights.	Authorization, Access Control	[32], [33], [39], [36]
Authorities can assign features to roles or concrete entities.	Authorization, Access Control	[40], [32], [29], [33], [36]
Each entity has access to the minimal amount of data possible.	Authorization, Access Control	[32], [29], [33], [38], [36]
There is an access timeout (or other control) set up and automatically disconnects the entity in case of inactivity.	Authorization, Access Control	[32], [33], [41], [36]
Authentication		
All (both local and remote) access should be protected by a unique entity identification and password, token, biometrics, or multi-factor authentication.	Access Control, Authentication	[32], [29], [33], [42], [37], [36]
Unsuccessful login attempts are logged and limited.	Access Control, Authentication, Logging	[32], [33], [36], [39], [44]
Used third-party technologies and components are certified that do not circumvent set IDs or passwords.	Access Control, Authentication, Third-party components	[32]
Used third-party configurations, updates, or other provided data use digital signatures.	Authentication, Third-party components	[32], [36]

Each guideline has one main phase and one main tag to be classified by. First-level classification is based on the incident phase, and is presented in Tables I–V. Inside these categories, guidelines are sorted by their main tag. This checklist view contains every guideline only once and is destined for the first walk-through to get to know all included recommendations. During incident planning, the architect should use it to make sure they reflect all the essential concerns and consider their properties. The phase-tags tree here may be a little more important than the recommendations themselves.

The architect should use the checklist either to guide the design process or at the end of the designing process to check that all the mentioned concerns are reflected. In fact, instead of ticking boxes, it is recommended to briefly *describe* the planned or realized extent to which the concern is reflected within the designed system. This can also improve communication within the development team or with stakeholders. Furthermore, when used to inspect the designed architecture, we recommend to annotate each concern with *strengths* and *weaknesses* of the designed solution with respect to the specific concern.

TABLE II
PHASE 2: DETECTION

GUIDELINE	TAGS	SOURCES
Logging		
All key events are logged.	Logging, Evidence	[46], [32], [47], [36]
Logs contain all important identification and classification.	Logging, Evidence	[32], [33], [36], [47]
Logs have various priority levels.	Logging	[32], [47]
There is a supervisory system that monitors logs and highlights alarms and anomalies.	Logging, Anomaly Detection	[32], [47], [48]
There is a mechanism that monitors if the logging system works.	Logging, Anomaly Detection	[32], [36]
Anomaly Detection		
The system recognizes distinct changes in configuration.	Anomaly Detection	[34], [29], [36]
The system recognizes unexpected or incomplete resets.	Anomaly Detection	[34]
The system recognizes memory failures.	Anomaly Detection	[34]
The system recognizes suspicious instructions.	Anomaly Detection	[34]
Anomalies in system performance are recognized and reacted to.	Anomaly Detection	[46], [48]
Anomalies in process behavior are recognized and reacted to.	Anomaly Detection	[46], [48]
File and directory changes are recognized and reacted to.	Anomaly Detection, Data Protection	[36], [46]
Input and Output		
User inputs and commands are validated and tested for sanity.	Input and Output	[49], [35]
External data are validated on entry.	Input and Output, Authentication, Third-party Components	[35]
The system controls all data before processing.	Input and Output	[34]
Phenomenon Evaluation		
Fault severity is classified into multiple levels.	Phenomenon Evaluation, Logging	[36], [48]
The system has a precisely defined failure tolerance threshold.	Phenomenon Evaluation	[32], [48]
Before launching a critical mode, the system checks if the trigger is valid.	Phenomenon Evaluation, Safety Ensuring	[32], [48]
Network		
Network data are collected.	Network, Evidence	[46]
All network alerts and error reports are checked.	Network, Anomaly Detection	[46]
The system recognizes unexpected, unusual, or suspicious traffic.	Network, Anomaly Detection	[46], [37]
Unauthorized entities connected to the system's network are recognized and restricted.	Network, Anomaly Detection, Authorization, Third-party Components	[29], [33], [36], [46]

TABLE III
PHASE 3: CONTAINMENT

GUIDELINE	TAGS	SOURCES
Stopping from Propagating		
The system is divided into independent parts with the possibility of a partial shutdown.	Stopping from Propagating	[34]
Safety-critical functions are isolated from non-safety-critical.	Stopping from Propagating	[46]
The system uses sandboxing to encapsulate high-risk parts.	Stopping from Propagating	[50]
Safety Ensuring		
The system is tolerant of an unstable or missing power source.	Safety Ensuring	[34]
Safety-critical software requirements are precisely identified and described.	Safety Ensuring	[30], [46]
There are *must-work functions* identified within the system.	Safety Ensuring, Self-Adaptiveness	[34], [51]
Must-work functions are redundant.	Safety Ensuring, Self-Adaptiveness	[34], [36], [51], [52]
Each of the *must-work functions* has at least two independent ways to control them.	Safety Ensuring	[34], [36]

TABLE IV
PHASE 4: RECOVERY

GUIDELINE	TAGS	SOURCES
Backups		
Critical data have a backup.	Backups	[29], [37]
Backups are off-site to be protected from local disasters (e.g., fire, flood, ...).	Backups	[36], [53]
Backup time intervals vary based on the frequency of changes.	Backups	[53]
Backups are protected from unauthorized access.	Backups, Access Control	[53]
Every backup process is checked to see if it was successful.	Backups	[29]
Before recovery from backup, data are preserved for further analysis of the incident.	Backups, Post-Incident Analysis	[29]
Self-Adaptiveness		
Must-work functions have a self-healing mechanism.	Self-Adaptiveness	[51]
System is prepared to adapt to operating without damaged parts.	Self-Adaptiveness, Safety Ensuring	[51]

B. Tags

Tags serve for a more detailed classification of the guidelines. Overall, they are based on system procedures, features, possibly used technologies, or specify parts of the incident phase more accurately. The tags, selected based on our methodology in Section IV, were also validated against the 20 common security requirements defined by CIPSEC [15]. They cover all these requirements except one requirement (number 15) that is management-oriented and thus out of the checklist scope, and add some requirements that were not covered by CIPSEC.

TABLE V
PHASE 5: POST-INCIDENT ANALYSIS

GUIDELINE	TAGS	SOURCES
Logging		
Logs are archived.	Logging, Evidence	[46]
Logs are secured.	Logging, Encryption	[46]
Old or useless logs are disposed of.	Logging	[46]
Evidence		
All possible evidence sources are identified.	Evidence	[54]
Evidence contains all necessary information to be classified as complete and valid.	Evidence	[33], [54]
All evidence data are secured.	Evidence, Encryption	[56], [57]
Evidence storage is reliable.	Evidence	[55]

a) Access Control: Access control ensures that data cannot be changed or read by unauthorized entities. This enhances their confidentiality and integrity, which are essential for software security [64].

b) Authentication: Authentication is a process of dedication and confirmation of the true identity of an external entity [42]. CI systems work with sensitive data, and their functionalities are safety-critical; therefore, access to them should be limited to trusted people and components.

c) Authorization: Authorization is deciding if a concrete authenticated entity is allowed to execute or make a specific action. Users and devices should have various roles based on their permission and have minimal possible rights [65], [39].

d) Data Protection: Data are often the main target of cyberattacks. They have to be protected from theft, unauthorized changes, or corruption [37].

e) Logging: Logging is the fundamental method to control and collect information about the program's behavior. It can help us detect an ongoing attack, gather evidence, improve the development process, etc. Therefore the whole mechanism is quite complex and should not be underestimated [47], [66].

f) Anomaly Detection: Anomaly detection means searching for deviations from the expected behavior of the system. All found anomalies should be inspected because they could signify an incident or danger. Correct detection can be tricky; possible false positives and negatives can be disastrous [48].

g) Input and Output (I/O): I/O vulnerabilities are not only frequent targets of attacks but also weak points vulnerable to inadequate usage by ordinary users. Operating a system without proper I/O handling is like leaving a house with doors and windows open [49].

h) Phenomenon Evaluation: This category covers a thin layer between anomaly detection and reaction to the incident. The reaction should not be reckless and hasty and must be evaluated adequately to the severity of the detected anomaly.

i) Network: With availability improvement and new technologies it is not imaginable to operate a CI system without any network connection anymore [67]. That brings up a variety of potential problems and vulnerabilities [68].

j) Stopping from Propagating: There can be many weak points across the software, and attackers may want to gain control of the entire system from them. Therefore we have to stop the spreading of any incident so it will not be possible for a localised fault to endanger distant parts of the system.

k) Safety Ensuring: Safety is an essential property of CI systems. As it strongly depends on the CI domain, we provide only an abstract solution to its ensuring within the system. This tag also indicates guidelines in other categories that may affect system safety.

l) Backups: System recovery depends directly on backups. Without proper backup, we may not be able to repair the system, and data will be lost. We must plan backups while designing the system because it would be too late to save data when an incident is detected. We also have to protect backups so they will not be damaged together with the system during the incident [69].

m) Evidence: Evidence is a cornerstone of forensic readiness. Identifying and collecting pieces of evidence should be taken into account already during the design of the system, not during the incident [70]. Evidence does not come by itself; we must be prepared to collect it and maximize its quality to facilitate the investigation [54], i.e. to ensure its integrity, prevent leaks, and protect contained sensitive data.

n) Third-Party Components: Third-party components often do not have as strict quality requirements as CI systems. Due to this, they may have safety issues that can propagate to our system. Therefore third-party components should be observed and not trusted by default [71].

o) Encryption: Encryption helps us preserve the integrity and confidentiality of the data within the system [72]. Similar to Network, this tag is mainly intended to identify encryption-related guidelines. Choice of concrete encryption techniques is out of the scope of architectural considerations.

p) Self-Adaptivness: CI systems may be too complex to be managed entirely by humans. Self-adaptive software monitors itself and its environment and reacts appropriately to detected changes. Therefore, such a system will be easier to maintain, and its responses to incidents will be faster [73].

VI. Conclusion

In this paper, we have presented a vision of an integrated checklist guiding the design of critical software systems, and presented first version of such a checklist. To this end, we did research to collect relevant standards and sources, all covering the scope only partially, and proposed a set of guidelines in the form of a checklist to enhance its straightforward usability during the software design. Guidelines were classified and sorted out to meet our defined checklist scope. Additionally, supplementary material is available at [7], containing the full guidelines classification data, detailed guidelines descriptions, and a demonstration of the checklist in a real-life context. In the future, we would like to validate the checklist with industrial experts and refining it with further views and layers of detail.

References

[1] I. Meedeniya, A. Aleti, and B. Buhnova, "Redundancy allocation in automotive systems using multi-objective optimisation," in *Symposium of Avionics/Automotive Systems Engineering (SAASE'09), San Diego, CA*, 2009.

[2] S. Chren, B. Rossi, B. Bühnova, and T. Pitner, "Reliability data for smart grids: Where the real data can be found," in *2018 smart city symposium prague (scsp)*. IEEE, 2018, pp. 1–6.

[3] J. Rodríguez, A. Galán, A. Alvarez, R. Díaz, and C. Consortium, *D2.1, CIPSEC System Design WP 2, Development of the CIPSEC security framework for Critical Infrastructure environments CIPSEC Enhancing Critical Infrastructure Protection with innovative SECurity framework*, Jan 2017.

[4] E. The European Union Agency for Cybersecurity, "Eu cybersecurity certification framework," Dec 2020. [Online]. Available: https://www.enisa.europa.eu/topics/standards/certification

[5] L. Bass, P. Clements, and R. Kazman, *Software architecture in practice, 3rd edition*. Addison-Wesley Professional, 2013.

[6] E. Fernandez-Buglioni, *Security patterns in practice: designing secure architectures using software patterns*. John Wiley & Sons, 2013.

[7] A. Bierska, B. Buhnova, and H. Bangui, "Supplementary material for the integrated checklist," https://drive.google.com/drive/folders/1DNjQdBmVTR7Z_JYZYpQS_QaohdFboIkC?usp=sharing, 2022.

[8] K. Lukitsch, M. Müller, and C. Stahlhut, "Criticality," in *Key Concepts for Critical Infrastructure Research*. Springer, 2018, pp. 11–20.

[9] "Federal ministry of the interior, national strategy for critical infrastructure protection, berlin, germany (www.bmi.bund.de, 2009."

[10] N. Medvidovic and R. N. Taylor, *Software architecture: foundations, theory, and practice*. John Wiley & Sons, 2010.

[11] G. Fairbanks, *Just enough software architecture: a risk-driven approach*. Marshall & Brainerd, 2010.

[12] M. Fowler, *Patterns of enterprise application architecture*. Addison-Wesley Longman Publishing Co., Inc., 2002.

[13] I. Jacobson, "The immense power of simple check-lists for monitoring projects," https://www.ivarjacobson.com/publications/blog/power-checklists, 2020.

[14] Z. A. Baig, "Multi-agent systems for protecting critical infrastructures: A survey," *Journal of Network and Computer Applications*, vol. 35, no. 3, pp. 1151–1161, 2012.

[15] C. Consortium, "D1.3, report on taxonomy of the ci environments," Feb 2018. [Online]. Available: https://www.cipsec.eu/sites/default/files/cipsec/public/content-files/deliverables/D1.3%20Report%20on%20Taxonomy%20of%20the%20CI%20environments.pdf

[16] B. Buhnova, T. Kazickova, M. Ge, L. Walletzky, F. Caputo, and L. Carrubbo, "A cross-domain landscape of ict services in smart cities," in *Artificial Intelligence, Machine Learning, and Optimization Tools for Smart Cities*. Springer, 2022, pp. 63–95.

[17] B. Robert, R. De Calan, and L. Morabito, "Modelling interdependencies among critical infrastructures," *International Journal of Critical Infrastructures*, vol. 4, no. 4, pp. 392–408, 2008.

[18] P. Cichonski, T. Millar, T. Grance, and K. Scarfone, "Computer security incident handling guide : Recommendations of the national institute of standards and technology," *Computer Security Incident Handling Guide*, vol. 2, Aug 2012. doi: 10.6028/nist.sp.800-61r2. [Online]. Available: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-61r2.pdf

[19] C. M. Machuca, S. Secci, P. Vizarreta, F. Kuipers, A. Gouglidis, D. Hutchison, S. Jouet, D. Pezaros, A. Elmokashfi, P. Heegaard *et al.*, "Technology-related disasters: A survey towards disaster-resilient software defined networks," in *2016 8th International Workshop on Resilient Networks Design and Modeling (RNDM)*. IEEE, 2016, pp. 35–42.

[20] S. S. Murtaza, W. Khreich, A. Hamou-Lhadj, and A. B. Bener, "Mining trends and patterns of software vulnerabilities," *Journal of Systems and Software*, vol. 117, pp. 218–228, 2016.

[21] F. Kadri, B. Birregah, and E. Châtelet, "The impact of natural disasters on critical infrastructures: A domino effect-based study," *Journal of Homeland Security and Emergency Management*, vol. 11, no. 2, pp. 217–241, 2014.

[22] G. Kirov, P. Zlateva, and D. Velev, "Software architecture for rapid development of hla-integrated simulations for critical infrastructure elements under natural disasters," *International Journal of Innovation, Management and Technology*, vol. 6, no. 4, p. 244, 2015.

[23] L. N. Alrawi and T. Pusatli, "Investigating end user errors in oil and gas critical control systems," in *Proceedings of the 2020 6th International Conference on Computer and Technology Applications*, 2020, pp. 41–45.

[24] T. Plėta, M. Tvaronavičienė, S. D. Casa, and K. Agafonov, "Cyber-attacks to critical energy infrastructure and management issues: Overview of selected cases," 2020.

[25] T. Limba, T. Plėta, K. Agafonov, and M. Damkus, "Cyber security management model for critical infrastructure," 2019.

[26] P. J. G. Seoane, "Use and limitations of checklists. other strategies for audits and inspections," *The Quality Assurance Journal: The Quality Assurance Journal for Pharmaceutical, Health and Environmental Professionals*, vol. 5, no. 3, pp. 133–136, 2001.

[27] M. Sibbald, A. B. de Bruin, and J. J. van Merrienboer, "Checklists improve experts' diagnostic decisions," *Medical education*, vol. 47, no. 3, pp. 301–308, 2013.

[28] E. Verdaasdonk, L. Stassen, P. P. Widhiasmara, and J. Dankelman, "Requirements for the design and implementation of checklists for surgical processes," *Surgical endoscopy*, vol. 23, no. 4, pp. 715–726, 2009.

[29] North American Electric Reliability Corporation - NERC, "Critical infrastructure protection standards," 2011.

[30] "IEEE standard for software safety plans," *IEEE Std 1228-1994*, pp. 1–24, 1993. doi: 10.1109/IEEESTD.1993.9097571

[31] D. G. Photovoltaics and E. Storage, "IEEE standard for interconnection and interoperability of distributed energy resources with associated electric power systems interfaces," *IEEE Std*, pp. 1547–2018, 2018.

[32] "IEEE standard for intelligent electronic devices cyber security capabilities," *IEEE Std 1686-2013 (Revision of IEEE Std 1686-2007)*, pp. 1–29, 2014. doi: 10.1109/IEEESTD.2014.6704702

[33] "IEEE standard cybersecurity requirements for substation automation, protection, and control systems," *IEEE Std C37.240-2014*, pp. 1–38, 2015. doi: 10.1109/IEEESTD.2015.7024885

[34] N. Aeronautics and S. Administration, "Nasa-std-8719.13 software safety standard," 2020.

[35] National Aeronautics and Space Administration, "Nasa-std-8739.8 software assurance and software safety standard," 2020.

[36] N. I. of Standards and Technology, "Nistir 7628 revision 1 – guidelines for smart grid cybersecurity," *The Smart Grid Interoperability Panel – Smart Grid Cybersecurity Committee*, 2014.

[37] O. Tayan, "Concepts and tools for protecting sensitive data in the it industry: a review of trends, challenges and mechanisms for data-protection," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 2, pp. 46–52, 2017.

[38] M. Papaioannou, M. Karageorgou, G. Mantas, V. Sucasas, I. Essop, J. Rodriguez, and D. Lymberopoulos, "A survey on security threats and countermeasures in internet of medical things (iomt)," *Transactions on Emerging Telecommunications Technologies*, p. e4049, 2020.

[39] B. W. Lampson, "Computer security in the real world," *Computer*, vol. 37, no. 6, pp. 37–46, 2004.

[40] E. B. Fernandez and J. Hawkins, "Determining role rights from use cases," in *Proceedings of the second ACM workshop on Role-based access control*, 1997, pp. 121–125.

[41] S. Mare, A. M. Markham, C. Cornelius, R. Peterson, and D. Kotz, "Zebra: Zero-effort bilateral recurring authentication," in *2014 IEEE Symposium on Security and Privacy*. IEEE, 2014, pp. 705–720.

[42] T. Nandy, M. Y. I. B. Idris, R. M. Noor, L. M. Kiah, L. S. Lun, N. B. A. Juma'at, I. Ahmedy, N. A. Ghani, and S. Bhattacharyya, "Review on security of internet of things authentication mechanism," *IEEE Access*, vol. 7, pp. 151 054–151 089, 2019.

[43] A. Naiakshina, A. Danilova, C. Tiefenau, M. Herzog, S. Dechand, and M. Smith, "Why do developers get password storage wrong? a qualitative usability study," in *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, 2017, pp. 311–328.

[44] M. Alsaleh, M. Mannan, and P. C. Van Oorschot, "Revisiting defenses against large-scale online password guessing attacks," *IEEE Transactions on dependable and secure computing*, vol. 9, no. 1, pp. 128–141, 2011.

[45] J. R. de Almeida, J. B. Camargo, B. A. Basseto, and S. M. Paz, "Best practices in code inspection for safety-critical software," *IEEE software*, vol. 20, no. 3, pp. 56–63, 2003.

[46] M. E. Whitman and H. J. Mattord, *Principles of incident response and disaster recovery*. Cengage Learning, 2021.

[47] A. Pecchia, M. Cinque, G. Carrozza, and D. Cotroneo, "Industry practices and event logging: Assessment of a critical software development process," in *2015 IEEE/ACM 37th IEEE International Conference on Software Engineering*, vol. 2. IEEE, 2015, pp. 169–178.

[48] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM computing surveys (CSUR)*, vol. 41, no. 3, pp. 1–58, 2009.

[49] J. K. Teto, R. Bearden, and D. C.-T. Lo, "The impact of defensive programming on i/o cybersecurity attacks," in *Proceedings of the SouthEast Conference*, 2017, pp. 102–111.

[50] M. Maass, A. Sales, B. Chung, and J. Sunshine, "A systematic analysis of the science of sandboxing," *PeerJ Computer Science*, vol. 2, p. e43, 2016.

[51] A. Tarinejad, H. Izadkhah, M. M. Ardakani, and K. Mirzaie, "Metrics for assessing reliability of self-healing software systems," *Computers & Electrical Engineering*, vol. 90, p. 106952, 2021.

[52] A. Mattavelli, "Software redundancy: what, where, how," Ph.D. dissertation, Università della Svizzera italiana, 2016.

[53] E. Nemeth, G. Snyder, S. Seebass, and T. Hein, *UNIX system administration handbook*. Pearson Education, 2000.

[54] R. Rowlingson *et al.*, "A ten step process for forensic readiness," *International Journal of Digital Evidence*, vol. 2, no. 3, pp. 1–28, 2004.

[55] L. Pasquale, D. Alrajeh, C. Peersman, T. Tun, B. Nuseibeh, and A. Rashid, "Towards forensic-ready software systems," in *2018 IEEE/ACM 40th International Conference on Software Engineering: New Ideas and Emerging Technologies Results (ICSE-NIER)*. IEEE, 2018, pp. 9–12.

[56] J. McQuaid, "Forensic considerations for cloud data storage - forensic focus," 2021. [Online]. Available: https://www.forensicfocus.com/webinars/forensic-considerations-for-cloud-data-storage/

[57] A. Singh, R. A. Ikuesan, and H. Venter, "Secure storage model for digital forensic readiness," *IEEE Access*, vol. 10, pp. 19 469–19 480, 2022.

[58] M. Hollick and S. Katzenbeisser, "Resilient critical infrastructures," in *Information Technology for Peace and Security*. Springer, 2019, pp. 305–318.

[59] M.-D. McLaughlin and J. Gogan, "Challenges and best practices in information security management," *MIS Quarterly Executive*, vol. 17, no. 3, p. 12, 2018.

[60] F. C. Freiling and B. Schwittay, "A common process model for incident response and computer forensics," *IMF 2007: IT-Incident Management & IT-Forensics*, 2007.

[61] E. C. Thompson, *Cybersecurity incident response: How to contain, eradicate, and recover from incidents*. Apress, 2018.

[62] R. Fateman, "Software fault prevention by language choice: Why c is not my favorite language," in *Advances in Computers*. Elsevier, 2002, vol. 56, pp. 167–188.

[63] A. Valjarevic and H. S. Venter, "Harmonised digital forensic investigation process model," in *2012 Information Security for South Africa*. IEEE, 2012, pp. 1–10.

[64] Q. He and A. I. Antón, "Requirements-based access control analysis and policy specification (recaps)," *Information and Software Technology*, vol. 51, no. 6, pp. 993–1009, 2009.

[65] K. Walsh, "Authorization and trust in software systems," 2012.

[66] G. Rong, Q. Zhang, X. Liu, and S. Gu, "A systematic review of logging practice in software engineering," in *2017 24th Asia-Pacific Software Engineering Conference (APSEC)*. IEEE, 2017, pp. 534–539.

[67] L. A. Maglaras, K.-H. Kim, H. Janicke, M. A. Ferrag, S. Rallis, P. Fragkou, A. Maglaras, and T. J. Cruz, "Cyber security of critical infrastructures," *Ict Express*, vol. 4, no. 1, pp. 42–45, 2018.

[68] G. Tzokatziou, L. Maglaras, and H. Janicke, "Insecure by design: Using human interface devices to exploit scada systems," in *3rd International Symposium for ICS & SCADA Cyber Security Research 2015 (ICS-CSR 2015) 3*, 2015, pp. 103–106.

[69] M. M. Howell, "Data backups and disaster recovery planning," 2003.

[70] L. Daubner, M. Macak, B. Buhnova, and T. Pitner, "Verification of forensic readiness in software development: A roadmap," in *Proceedings of the 35th Annual ACM Symposium on Applied Computing*, 2020, pp. 1658–1661.

[71] D. E. Rico and M. Hann, "A combined dependability and security approach for third party software in space systems," *arXiv preprint arXiv:1608.06133*, 2016.

[72] J. Obert, P. Cordeiro, J. T. Johnson, G. Lum, T. Tansy, N. Pala, and R. Ih, "Recommendations for trust and encryption in der interoperability standards," Sandia National Lab.(SNL-NM), Albuquerque, NM (United States); Kitu Systems, Tech. Rep., 2019.

[73] M. Salehie and L. Tahvildari, "Self-adaptive software: Landscape and research challenges," *ACM transactions on autonomous and adaptive systems (TAAS)*, vol. 4, no. 2, pp. 1–42, 2009.

Author Index