POLITECHNIKA LUBELSKA

MONOGRAFIE

Computational Intelligence, Information Systems and Data Mining

edited by
Małgorzata Charytanowicz
Paweł Karczmarek
Adam Kiersztyn

Lublin 2021

Computational Intelligence, Information Systems and Data Mining

Monografie – Politechnika Lubelska

Politechnika Lubelska
Wydział Elektrotechniki i Informatyki
ul. Nadbystrzycka 38A
20-618 Lublin

Computational Intelligence, Information Systems and Data Mining

edited by
Małgorzata Charytanowicz
Paweł Karczmarek
Adam Kiersztyn

Wydawnictwo
Politechniki Lubelskiej

Lublin 2021

Reviewers:
Piotr A. Kowalski
Grzegorz Kozieł
Edyta Łukasik
Dorota Pylak

Publication approved by the Rector of Lublin University of Technology

ISBN: 978-83-7947-492-9

Publisher: Wydawnictwo Politechniki Lubelskiej
www.biblioteka.pollub.pl/wydawnictwa
ul. Nadbystrzycka 36C, 20-618 Lublin
tel. (81) 538-46-59

Printed by: Soft Vision Mariusz Rajski
www.printone.pl

The digital version is available at the Digital Library of Lublin University of Technology: www.bc.pollub.pl

Contents

Part I: Artificial Intelligence and Information Technology

Peri Bakasova, Nella Israilova
Chat bot – Virtual Teacher Assistant 9

Zainelkhriet Murzabekov, Marek Milosz, Kamshat Tussupova
The finite horizon optimal control problem for a nonlinear model of a three-sector economic 24

Dauren Nazarbayev, Zhanna Alimzhanova
Adaptation of dynamic models to cryptanalysis conditions 35

Diana Rakhimova, Aliya Turganbayeva, Akbota Kulzhanova, Alima Suleimenova
Semantic analysis of the Kazakh language based on machine learning 48

Kunduz T. Sharsheeva, Gulnaz U. Tultemirova
The study of methods for optimizing the route of movement in the city 59

Gulnaz Tultemirova, Kunduz Sharsheeva, Gulnara Oruzbaeva, Abdyldabek Akkozov
Computer model of the holograms synthesis with a real phase 73

Salamat Zhunusbayeva, Zhanna Alimzhanova
Use of new technologies in cryptanalysis 84

Part II: Data Analysis and Decision Making

Katarzyna Baran
Thermal imaging of stress: a review 95

Szymon Fornal, Paweł Karczmarek
Application of the AHP method to analyze the state of knowledge of students in particular years of studying 114

Adrian Jaczyński, Paweł Karczmarek
Analysis of the use of e-learning platforms in connection with the COVID-19 virus 137

Adam Kiersztyn, Krystyna Kiersztyn, Patrycja Jędrzejewska-Rzezak, Paweł Karczmarek, Witold Pedrycz
Analysis of self-awareness of beginning programmers 153

Adam Kiersztyn, Pavel Urbanovich, Nadzeya Shutko
The concept of random cluster based outlier detection 170

Michał Kuśpit, Paweł Karczmarek
A comprehensive experimental comparison of COVID-19 epidemiological models 182

Preface

Nowadays, information technology and computational intelligence are dynamically developing fields of science that cover not only areas traditionally assigned to mathematics and computer science, but also to other sciences dealing with collecting and processing information using computer techniques.

The monograph “Computational Intelligence, Information Systems and Data Mining” presents scientific investigations and research directions conducted at the Department of Computer Science at the Lublin University of Technology and cooperating Universities. This aims to inform the reader of new theoretical approaches and methodologies, as well as state of the art information technology and its applications. The research activities focus on many fields of computer science, including artificial intelligence, intelligent data analysis methods, machine learning, computational linguistics, machine vision, information security, decision support systems and others. The articles contained in this publication are the result of in-depth knowledge of modern IT tools supporting decision-making processes and presenting various approaches used in data analysis.

All the Authors are hopeful that their research works will be interesting to a wide global readership.

We would like to express our gratitude to the Authors who contributed their original research papers, as well as to all reviewers for their valuable comments.

August 2021

Editors
Małgorzata Charytanowicz
Paweł Karczmarek
Adam Kiersztyn

Part I

Artificial Intelligence and Information Technology

Peri Bakasova[1], Nella Israilova[2]

Chat bot – Virtual Teacher Assistant

Abstract: The coronavirus pandemic has forced educational institutions to switch to distance learning. Despite the fact that not everyone turned out to be ready for online learning, the teachers switched to the distance format. To organize online learning, teachers and students learned to use educational portals, various online services, video conferencing, instant messengers and YouTube. In the context of distance learning, the participants faced various difficulties in the form of a lack of electricity, Internet and other force majeure. But the main problem was to ensure fast communication between the teacher and the student. If they have questions outside the classroom, students turn to the teachers in messengers, but they do not always receive a timely response due to the teacher's employment. The question loses its relevance and the problem remains unsolved. The repetition of such cases can lead to a complete loss of communication between the student and the teacher, resulting in a decrease in interest in the studied discipline. Accordingly, teachers need to take action to quickly respond to student questions and help with problem solving. Students' questions are often identical: request to send links to videoconferences, class materials, etc.; registration in the online service; warning about possible absence from class. These requests can be passed to artificial intelligence: chat bot can be used to answer standard questions (according to the FAQ principle, but we understand the text in NL and user requests), for routing the user to the required section of the course where he can find materials, and also as "prompts" – bots for intelligent prompts to the teacher. All this allows you to significantly relieve the teacher. However, the AI + Human link is most effective, when unique questions are transferred to the teacher, and he devotes a sufficient amount of time to the student to help solve the problem. This article explores the development of chat bot as a virtual teaching assistant.

Keywords: AI, artificial intelligence, NL, natural language, chat bot, development, online learning, NLU, ML, NLP

1. Introduction

In the age of global digitalization, new technologies and services are being actively created and implemented that are effectively used in the educational process and that meet the needs of modern students. In the past few years, and more actively during the pandemic, the current trend is the creation of chat bots, which have enormous potential and in the near future can replace many applications and services.

By definition, a chat bot is a computer program that can "communicate" with a person in a common language in text or voice form, through an intuitive interface.

In modern education, the relevance of creating chat bots for solving various kinds of problems within educational organizations is growing rapidly; many

[1] Peri Bakasova, KSTU named after I.Razzakov, Bishkek, Kyrgyzstan;
e-mail: bakasovap@mail.ru
[2] Nella Israilova, KSTU named after I.Razzakov, Bishkek, Kyrgyzstan;
e-mail: inela.kstu@gmail.com

universities in the world are already using such digital assistants. Chat bots in educational organizations are focused on solving various problems and have a wide range of functions. If we consider the functionality of chat bots in the educational process, then it can solve both simple organizational issues and more complex ones, performing the functions of a teacher or student assistant. For example, depending on the context, chat bots can be used for the following purposes:

- distribution and control of the implementation of practical tasks, information support (step-by-step tips, leading questions, etc.);
- analysis of the text for errors with the output of a set of recommendations;
- feedback in real time with answers to typical questions of each student, freeing up the teachers' time for qualified activities;
- organizing the collection of information and analysis of student behavior to build an individual educational trajectory;
- testing and verification of learning outcomes for a set of adaptive parameters.

The attractiveness of chat bots is explained by the fact that they provide at least: anonymity, efficiency, the ability to maintain dialogue, and personalization. Among the advantages in favor of introducing chat bots, one can also note the fact that they are easy to install, do not take up device memory resources, and are convenient for distribution.

In this article, a chat bot is considered as a teacher's assistant, which allows the teacher to reduce the time spent on organizing the course and provides operational interaction with students. Saved time can be successfully invested by the teacher in research work and in guiding and motivating the study group.

The chat bot provides the student with personalized learning, by adapting to the individual pace and rhythm of the student, in accordance with his individual needs and requirements.

Thus, the chat bot is a very useful tool in organizing the educational process, as well as an interesting and easy-to-use tool for both learners and teachers in distance learning.

2. How chat bots work

In general, a typical work cycle of any chat bot can be represented as a chain of the following actions:

- receiving a request from the user;
- parsing the request – understanding the statement and determining the user's intentions;
- execution of actions according to a predetermined scenario for processing a user case;
- generating a response in natural language;
- saving the request, context and parameters of the dialogue for processing subsequent calls;
- sending a response to the user.

The most difficult stage of the work is parsing the client request. Until about 2015, when developing chat bots, the approach based on formal rules (rule-based) was mainly used. Its essence consists in highlighting semantically significant elements of phrases, their codification, creating special formal script programming languages that allow describing dialog scenarios (e.g. Javascript, PHP, Python). However, after 2015, the development of algorithms for semantic proximity of texts, technologies for speech synthesis and recognition, as well as Big Data and Machine Learning, has led to the spread of new approaches to text classification and training of natural language understanding systems (NLU, Natural Language Understanding). Thus, most modern chat bots are based on the latest advances in data science: NLU and NLP technology, speech recognition and text processing methods using neural networks and other artificial intelligence tools.

Natural language understanding (NLU) is a branch of artificial intelligence (AI) that uses computer software to understand input made in the form of sentences in text or speech format.

NLU directly enables human-computer interaction (HCI). NLU understanding of natural human languages enables computers to understand commands without the formalized syntax of computer languages and for computers to communicate back to humans in their own languages.

The field of NLU is an important and challenging subset of natural language processing (NLP). While both understand human language, NLU is tasked with communicating with untrained individuals and understanding their intent, meaning that NLU goes beyond understanding words and interprets meaning. NLU is even programmed with the ability to understand meaning in spite of common human errors like mispronunciations or transposed letters or words.

NLU uses algorithms to reduce human speech into a structured ontology. AI fishes out such things as intent, timing, locations and sentiments. For example, a request for materials of a lecture held on February 18 on the course System Programming might look something like this: Lecture materials [intent] / System Programming course [course name | location] / February 18 [date].

2.1. Stages of the process of “understanding” natural language

For modern text chatbots, the parsing process includes the following steps:

- text preprocessing:
  - tokenization (word splitting);
  - correction of typos;
  - lemmatization and stemming (determination of the normal form of words and parts of speech);
  - dropping stop words (articles, interjections, conjunctions, etc.);
  - expanding the query using dictionaries of synonyms;
  - supplementing information on the significance of individual words;

- expanding the query with a parse tree and pronoun resolution results;
- definition of named entities;

- request classification:
  - based on examples of phrases and ML-algorithms or formal rules (templates);
  - ranking the classification hypotheses according to the current context of the conversation;
- retrieving query parameters from a user's phrase.

3. How do conversational AI technologies work?

Initially, the user addresses his request to any of the channels available to him. The channels can be smart devices, assistants built into devices or mobile phones, as well as instant messengers.

In our case, we chose Telegram messenger for several reasons. Firstly, Telegram provides a lot of resources for creating chat bots, so it's easier and cheaper to make them. Secondly, the Bot in Telegram can be linked to the company's website or other services, set a list of commands to it and set up a menu for users. Thirdly, the ability to work at low Internet speeds. And the most important thing is that our students have chosen this messenger as the most convenient for sharing files.

So we got a user request. There is some intent behind the request, i.e. desire to receive an answer to a question, to receive a service or any content, for example, a link, lectures or videos.

Further, additional processing or conversion of the message format may be required. Dialogue platforms always work with text, while a number of channels involve voice communication. The ASR (speech recognition), TTS (speech synthesis) platforms, telephony integration systems are responsible for this conversion. In some cases, it may be necessary to recognize the interlocutor by voice – in this case, biometrics platforms are used. Separate channels, for example, messengers or Alice's assistant on a mobile phone, allow you to combine visual interactive elements (for example, buttons or product cards that you can tap on) and natural language. To work with them, you need to integrate with the corresponding APIs. In our development, we will skip this step and assume that the request came in text format.

The request, converted to text, goes to the dialog platform. Its task is to understand the meaning of what has been said, to capture the user intent and effectively process it, giving the result. Intentions indicate what the user wants, but discards information about how he wrote about it. For this, dialogue platforms use many technologies, such as text normalization, morphological analysis, analysis of the semantic proximity of what has been said, ranking hypotheses, highlighting named entities and, finally, generating queries in machine language through a set of APIs to external databases and information systems.

Having received the data, the dialogue platform generates a response – a text, a voice message (using TTS), turns on content streaming or notifies of a completed action (for example, provides access to course materials). If there is not enough data in the initial request for making decisions on further action, the NLU platform initiates a clarifying dialogue in order to get all the missing parameters and remove the uncertainty.

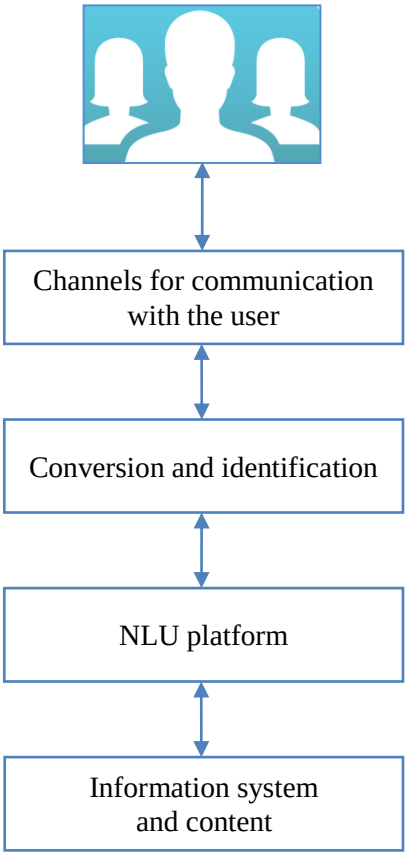


Fig. 1. Scheme of interaction between user and chat bot

4. The logic of working with the bot

We need to make a virtual assistant that helps the teacher and answers students' questions. To do this, you need to complete the following steps.

Stage 1

It is necessary to come up with possible user phrases and describe the bot's reaction to these phrases. As an example, suppose a student makes a request for

course materials. The basic query is clear – find files. But this request can have many parameters: discipline, number or topic of the lesson, date of the lesson, format of the material, etc. Perhaps the user will immediately write the name of the file to be found. Or he will clarify the details: “What is the size of the file?” etc. Moreover, in addition to inquiries about the case, there are other small talk phrases that need to be asked.

When it seems to us that the described situations will be enough to cover 80% of requests at best, we stop. Then we scatter our assumptions about user needs for specific requests – intentions.

List of intentions

Intent 1 *– Materials are needed*

Phrases:

- Can you poison the materials?
- Where can I find lecture materials?
- ...

What to do: randomly select one of the most recent lectures and send the user a link to the lecture materials.

Intent 2 *– Course materials are needed*

Phrases:

- lectures on SP are needed
- are there pdf lectures for this course?
- need another lecture of this course
- ...

What to do: add a course filter to your request and choose one of the most popular topics.

Intention 3 *– materials of the 1st lecture of the course are needed*

Phrases:

- is there a pdf of the first lecture on joint venture?
- Today we had a lecture on the joint venture. Can you send materials?
- ...

What to do: a query to the database by the criterion [topic | lecture number] && [course name] select materials and send to the user.

Intention N *– we misunderstood the user*

Phrases: any others

What to do: issue a dummy phrase: “I don’t know what you meant, but you can find the course materials on the department’s website.”

Stage 2

- At this stage, we start creating an algorithm according to which the robot will work.
- There are 2 ways:

- **Simple**: You need to manually define keywords for each intention, and when the user writes, the bot will search for these keywords in the phrase.
- More **complex**, but effective: Train the algorithm to compare the user's replicas by lexical meaning. This allows you to find the closest intention.

Table 1. Comparison of a keyword-based bot's reaction versus an AI bot's reaction at users query

Keyword-based bot	AI bot	Note
User: Hi! Please send today's lecture on the OS. **Bot**: Hello! **Bot**: What discipline? **User**: by OS **Bot**: link_to_materials_lections_date_course	**User**: Hi! Please send today's lecture on the OS. **Bot**: Hello! **Bot**: link_to_materials_lections_date_course	In the 1st case, the Bot – followed by keywords *send + today + lecture*, did not understand the name of the course and asked again. In the 2nd – bot went according to the closest intention and extracted the name of the course/location
User: I did not understand the condition of the 3rd task. **Bot**: link_to_3e_course_set	**User**: I did not understand the condition of the 3rd task. **Bot**: Detailed instructions for completing the 3rd task can be found at link_to_description_3rd_assignment	In the 1st case, the bot was guided by the keyword "condition" and was mistaken. In the 2nd case, the bot realized that the intention closest to the "condition did not understand" is to explain the condition
User: Submitted the 2nd task. The other 3 have already been handed over. **Bot**: link_to_3e_course_set	**User**: Submitted the 2nd task. The other 3 have already been handed over. **Bot**: Good! You have submitted all the work	In the 1st case, text processing occurs one sentence at a time. The bot saw "sent 2nd", responded using the template and offered the next task – in the end it didn't guess right and got confused. In the 2nd case, the general picture is formed

Consider algorithms for determining intentions for each method.

Table 2. Comparison of a keyword-based bot's versus an AI bot's algorithms

Keyword-based bot	AI bot
To understand the user, we take keywords. When creating a bot, machine learning and NLU technologies are not used	To understand the user, we make a smart comparison using Machine Learning and NLU algorithms and choose the closest intent
Preparation: For each intention we write out keywords (today's lectures on OS: today, lectures, files, OS, lectures + OS) **Algorithm:** • we take words from the user's replica ("Recommend wine to steak" → advise, wine, steak); • select the intention with the most keywords from the user's replica (wine_pod_meat); • if the bot did not find intersections by keywords, select the intention "we did not understand the user"; • we perform an action that corresponds to the intention	**Preparation:** Train the Model for Smart Comparison **Algorithm:** • compare the user's phrase with intent phrases using a smart comparison; • choosing an intention with the closest phrases; • if the selected intent is not close enough to the user's replica (the set threshold has not been passed), select the intent "we did not understand the user"; • we perform an action that corresponds to the intention
How to improve: To reduce errors, you need to add commands and buttons. Then users will stop communicating in the language altogether, but will simply click on the buttons. Such improvements lead to the degradation of spoken intelligence	**How to improve:** After users chat with the robot, we will have new examples of phrases and we will distribute them by intent. Over time, the percentage of coverage will increase, the assistant will begin to better cope with communication. These improvements lead to improved conversational intelligence

5. Request processing logic

The comparative table (Table 2) shows that the second method gives more correct results than the first. Therefore, when developing our bot, we used the second method. Now let's look at the logic for processing requests for our bot.

The general scheme of our platform operation can be represented as follows:

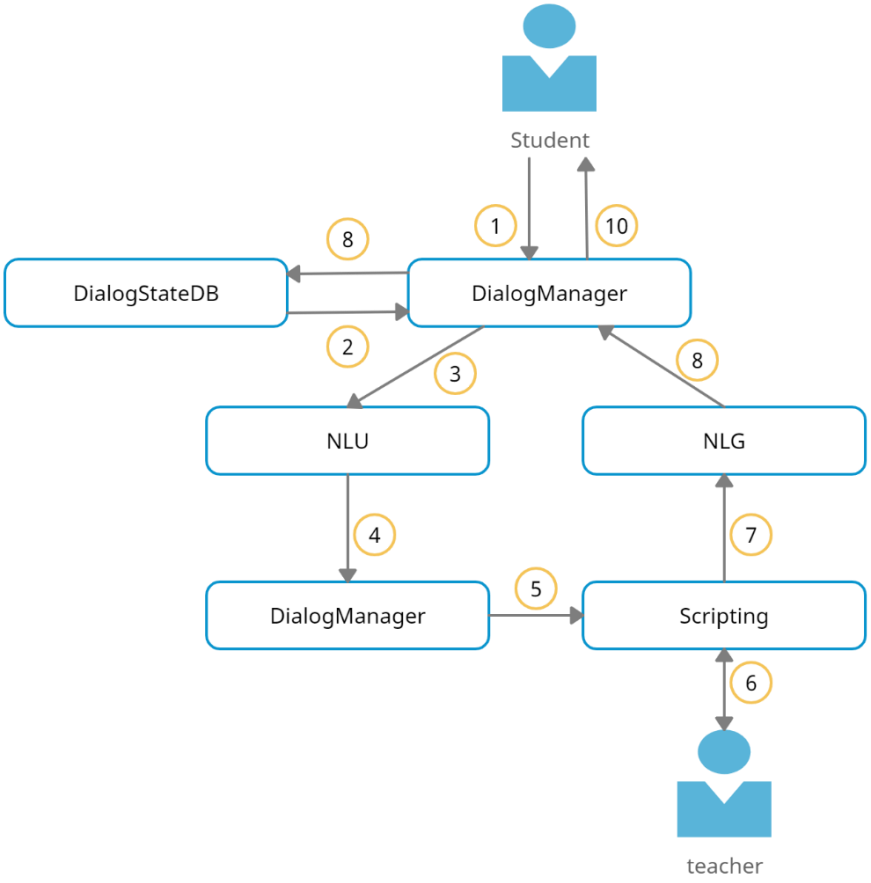


Fig. 2. The algorithm of the virtual assistant

The main loop for processing a student's request consists of the following events and actions:

- the dialog management module DialogManager accepts the client's request;
- DialogManager loads the dialog context from the database;
- the client's request is sent for processing to the NLU-module to determine the client's intention and its parameters;
- based on the dialog script and the extracted data, DialogManager determines the next state (block, screen, dialog page) that most closely matches the client's request;
- execution of business logic (scripts) in accordance with the given script of the chatbot;
- if the bot cannot cope on its own, then it sends a request to the teacher and expects a response;

- formation of a response in natural language according to a specific scenario or a response from a teacher;
- passing the response to DialogManager;
- saving the context and parameters of the dialogue for subsequent calls;
- sending a response to the user.

An important part of the system's operation is dialogue flow control (DialogManager), within which the overall message context and communication with previous and subsequent operators are defined. Thanks to this process, one or another phrase will be perceived differently, depending on when the message was received, who sent it, what additional data was transferred to the system along with the request. We propose to move these functions to the "script" level of the dialogue, so that this process is fully controlled by the bot developer.

Bot action scripts based on dialog scripts:

- FAQ dialogues;
- answers to typical questions of students in messengers (requests for the organization of the discipline, points, deadlines, laboratory assignments, deadlines for exams and laboratory work);
- directing the student to the sections of the course, where he can find materials on his request (literature, multimedia: audio and video files);
- if the bot does not cope with the dialogue on its own, then it sends a request to the teacher. He acts as a "prompter" – a bot for intellectual clues to the teacher.

Redirection can happen in the following cases:

- when the request may contain intentions, actions for which were not provided by the developer;
- the bot was unable to determine the intention even after clarifying questions.

As a prompter, the bot performs the following functions:

- redirects the student's questions to the teacher;
- offers options for answers or actions. In the event that the bot was able to determine the intent, but does not have appropriate indications for this intent. Popular answers will be listed as suggested answers:
- redirects the teacher's answers to the student;
- learns.

In general, according to the scenario, the chatbot should advise the student on its own, without the participation of the teacher. Those the bot must analyze the information received from the user, clarify it, generate and send a response. If the bot cannot cope with the dialogue, then the chat will be redirected to the teacher. In this case, the bot must also redirect the history of the student dialogue to the teacher. AI + Human package is the most efficient approach to bot implementation.

6. Intentions and actions

Above, we have listed the scenarios in which our bot can operate. Scenarios contain descriptions of all situations that can happen in a dialogue between a person and a machine. To describe the situation, you need to associate the names of intentions and actions.

```
user_query(['send materials of a lecture held on February 18 on the course System Programming'])
>>> [['todo: send materials', 'lesson_type: lecture', 'date: February 18', 'location: System Programming']]
```

The list of available tags and their descriptions are presented below.

Table 3. The list of available tags and their descriptions

Tags	Descriptions
person	person
course	course name
lesson_type	lectures, seminar, practice, laboratory exercises
organization	companies, agencies, institutions, etc.
gpe	countries, cities, states
location	site location
product	materials, files, documents, archives
language	any named language
date	absolute or relative dates or periods
time	times smaller than a day
percent	percentage (including “%”)
todo	action, verb

Let’s look at a simple example – a greeting:

```
if intent(привет) { reaction(сказать_привет) }
```

It looks like a regular code, but behind the intent (...) construction, a neural network is used to describe the dialogue in general patterns (“if you were asked for this or that”) using ordinary programming constructs.

```
if intent(hello) {
    if was_reaction(say_hello) {
          reaction(say_that_have_already_said_hello) }
    else {
          reaction(say_hello) }
}
```

It says here: if you said hello, say hello in return. And if after that they said “hello” again – say that you have already said hello.

A reaction is an action that a bot must perform in response to an intent. In 95% of cases, it is just text. But the robot can also call a function in code, switch communication to the teacher, or perform other complex actions.

The code for sending text and functions itself exists separately from the language – the language describes situations as simply as possible.

Consider another example, where users are asked to send materials by mail or chat.

```
if intent(send_files) or intent(send_files_to_chat)  {
      reaction(send_files_to_chat) {
           if intent(send_to_mail) {
                reaction(send_files_to_mail)
          }
      }
}
if intent(send_files_to_mail) {
    reaction(send_files_to_mail)
}
```

Two situations are described here:

- if asked to send details – send them to the chat. If after they asked “to the mail”, then send them to the mail;
- if you immediately asked to send the details to the mail – send the details to the mail.

This is how the bot accomplishes its task – it works in context. Even if a person writes “to the mail” the robot will understand that we are talking about details.

7. Conclusions

Using of chat bots in the educational process will allow students to receive quick answers as questions arise, and reduce the workload of teachers. The presence of feedback will allow you to identify unresolved issues and supplement the course materials.

In general, the chatbot is useful, interesting and easy to use for both students and teachers.

Training on real dialogues is necessary to increase the intelligence of the chatbot. The bot must accumulate a certain amount of dialogues, i.e. he must write down everything and keep a history of correspondence. On these dialogues, you can train the bot and improve the quality of recognition of words, endings, cases and different turns of speech.

Programming and launching your own chatbot based on machine learning is a complex process that requires qualified developers and interface specialists, as well as significant time and resource costs. However, it is now possible to create a simple chatbot using many services without any special skills and knowledge of programming languages.

Thus, the chatbot is a very useful tool in organizing the educational process, as well as interesting and convenient to use for both students and teachers. Among other things, it meets the needs of the younger generation who are gaining knowledge in the context of digitalization.

The bot considered in the article is adapted only for the messenger, but it is planned to develop the functionality and link the bot to the LMS Learning Management Systems.

Bibliography

[1] Shevat A., Designing Bots: Creating Conversational Experiences, 2017.

[2] Nass C., Wired for Speech: How Voice Activates and Advances the Human-Computer Relationship, 2005.

[3] Kaplan J., Artificial Intelligence: What Everyone Needs to Know, 2017.

[4] Golden K., The Best Interface Is No Interface: The simple path to brilliant technology, 2015.

[5] Howard C., Hapke H., Lane H., Natural Language Processing in Action – Manning Publications, 2019.

[6] Anfimova E.A., Generation Z: problems, opportunities, prospects on the labor market, Innovative economy: prospects for development and improvement, 2018, 7(33): 256–261.

[7] Katkalo V.S., Corporate training for the digital world: a tutorial, Katkalo V.S., Volkova D.L., 2nd ed., Rev. and add., 2018, 248.

[8] Okulov S.A., Formation of a management system for the educational process by means of information technology, Okulov S.A., Success of modern science, 2017, 5: 170–174.

[9] Provotar A.I., Features and problems of virtual communication using chat bots, Provotar A.I., Klochko K.A., Applied and Computational Linguistics, 2018, 3: 2–7.

[10] Firsova E.A., Prospects for the use of chat bots in higher education, Informatization of Science and Education, 2018, 3(35): 157–166.

[11] Automation of answers to frequently asked questions in the skill for “Alice” using the DeepPavlov library / Blog of the company Moscow Institute of Physics and Technology, Habr., https://habr.com/ru/company/mipt/blog/445748/. Last accessed: 29.02.2021.

[12] Conversational AI: how chat bots work and who makes them, Just AI company blog, Habr., https://habr.com/ru/company/just_ai/blog/364149/. Last accessed: 29.02.2021.

[13] 8 services for creating a chatbot. SendPulse Blog, https://sendpulse.com/ru/blog/chat-bot-services/. Last accessed: 29.02.2021.

[14] Academic Earth, https://academicearth.org/ (дата обращения: 29.06.2020). Last accessed: 02.02.2021.
[15] Alexa and Siri Can Hear This Hidden Command. You Can't, https://www.nytimes.com/2018/05/10/technology/alexa-siri-hidden-command-audio-attacks.html. Last accessed: 29.02.2021.
[16] Artificial intelligence trends (gartner.com), https://www.gartner.com/smarterwithgartner/top-trends-on-the-gartner-hype-cycle-for-artificial-intelligence-2019. Last accessed: 29.02.2021.
[17] Big Data, Kafka, Hadoop, Spark, Arenadata, NoSQL (bigdataschool.ru), https://www.bigdataschool.ru/. Last accessed: 29.02.2021.
[18] CCMatrix: A billion-scale bitext dataset for training translation models, https://ai.facebook.com/blog/ccmatrix-a-billion-scale-bitext-data-set-for-training-translation-models/. Last accessed: 29.02.2021.
[19] Chatbots: an introduction from the developer, https://proglib.io/p/chat-bots-intro/amp/. Last accessed: 29.02.2021.
[20] Codeacademy, https://www.codecademy.com/ (дата обращения: 29.06.2020). Last accessed: 29.02.2021.
[21] DeepPavlov., http://deeppavlov.ai/. Last accessed: 29.02.2021.
[22] Dialogflower – Google Dialogflow (habr.com), https://habr.com/ru/post/412863/. Last accessed: 15.02.2021.
[23] edX, https://www.edx.org/. Last accessed: 29.02.2021.
[24] ELIZA, https://en.wikipedia.org/wiki/ELIZA. Last accessed: 29.02.2021.
[25] Enterprise Artificial Intelligence information, news and tips – SearchEnter prise AI, https://searchenterprisea.techtarget.com/. Last accessed: 29.02.2021
[26] Facebook is shutting down M, its personal assistant service that combined humans and AI. The Verge. Last accessed: 29.02.2021.
[27] Gartner I., Virtual Customer Assistants Market Gartner review, https://www.gartner.com/reviews/market/virtual-customer-assistants. Last accessed: 29.02.2021.
[28] Insider B., Amazon, Apple, Google, IBM, and Microsoft worse at transcribing black people's voices than white people's with AI voice recognition, study finds, https://www.businessinsider.com/study-ai-voice-recognition-racially-biased-against-black-voices-2020-3. Last accessed: 29.02.2021.
[29] Just AI, https://just-ai.com/. Last accessed: 29.02.2021.
[30] Khan Academy, https://www.khanacademy.org/. Last accessed: 29.02.2021.
[31] Markets R.A., Intelligent Virtual Assistant (IVA) Market – Growth, Trends, and Forecast (2020–2025), https://www.researchandmarkets.com/reports/4845914/intelligent-virtual[1]assistant-iva-market. Last accessed: 29.02.2021.
[32] Memrise, https://www.memrise.com/. Last accessed: 29.02.2021.
[33] Named Entity Recognition (NER) – DeepPavlov 0.14.1 documentation, http://docs.deeppavlov.ai/en/master/features/models/ner.html. Last accessed: 29.02.2021.

[34] Quizlet, https://quizlet.com/. Last accessed: 29.02.2021.
[35] Udemy, https://www.udemy.com/. Last accessed: 29.02.2021.
[36] Woebot, https://www.woebot.io/#intro. Last accessed: 29.02.2021.
[37] Coursera, https://www.coursera.org/. Last accessed: 29.02.2021.
[38] Умные экраны с камерами, биометрия, контекстуальность и интерактивные драмы – как виртуальные ассистенты станут реальными, https://rb.ru/longread/virtual-assistants/. Last accessed: 29.02.2021.
[39] Яндекс. Навыки Алисы. https://dialogs.yandex.ru/store/. Last accessed: 29.02.2021.
[40] Enterprise Artificial Intelligence information, news and tips – SearchEnterprise AI. https://searchenterpriseai.techtarget.com/. Last accessed: 29.02.2021.
[41] Big Data, Kafka, Hadoop, Spark, Arenadata, NoSQL (bigdataschool.ru), https://www.bigdataschool.ru/. Last accessed: 29.02.2021.
[42] Automation of answers to frequently asked questions in the skill for “Alice” using the DeepPavlov library. Blog of the company Moscow Institute of Physics and Technology (MIPT). Habr., https://habr.com/ru/company/mipt/blog/445748/. Last accessed: 29.02.2021.

Zainelkhriet Murzabekov[1], Marek Milosz [2], Kamshat Tussupova [3]

The finite horizon optimal control problem for a nonlinear model of a three-sector economic

Abstract: The problem of optimal control over a finite time interval is posed for the mathematical model of a three-sector economic. In this paper, we consider the optimal control problem for a class of nonlinear systems with fixed ends of trajectories without restrictions and with constraints on controls. We propose an algorithm for solving the optimal control problem for a nonlinear system with a given quality functional. Nonlinear control based on the principle of feedback with respect to control constraints is found. The problem is solved using Lagrange multipliers of a special type, which allows to find synthesising control that depends on the system's state and the current time. The results obtained for nonlinear systems are used in the construction of control parameters for the mathematical model of a three-sector economic cluster over a finite time interval with a given functional and various initial conditions. The results of system state calculations are presented in figures, optimal controls satisfying specified limitations. Optimal distribution of labour and investment resources are determined for a case study considered in the paper. They ensure bringing the system to an equilibrium state and satisfy balance ratios.

Keywords: optimal control problem, three-sector economic cluster, Lagrange multipliers method, nonlinear system, quadratic cost functional

1. Introduction

The three-sector economic model and the necessary conditions for optimal balanced economic growth were given by Kolemayev [1] and Zhang [2]. The extensive work of Aseev et al. [3] provided a basis of the mathematical theory of the infinite-horizon optimal control of dynamical systems using the Pontryagin maximum principle and an example of a two-sector model of optimal economic growth with an occasional jump in prices.

It should be noted that the controllability criteria for nonlinear systems were obtained in Klamka [4], and for discrete systems in Klamka [5]. Miłosz et al [6] considered optimisation of discrete processes with bounded control. Stabilisation problems for linear systems described by ordinary differential equations and partial differential equations are discussed by Mitkowski et al. [7]. Afanas'ev and Orlov [8] consider a class of nonlinear systems for which there exists a coordinate representation (diffeomorphism) that transforms an original system into a system with a linear fundamental part and a nonlinear feedback.

[1] Zainelkhriet Murzabekov, Al-Farabi Kazakh National University, ave. al-Farabi 71, Almaty, Kazakhstan; e-mail: murzabekov-zein@mail.ru

[2] Marek Milosz, Department of Computer Science, Lublin University of Technology, Nadbystrzycka 36B, 20-618 Lublin, Poland; e-mail: m.milosz@pollub.pl

[3] Kamshat Tussupova, Al-Farabi Kazakh National University, ave. al-Farabi 71, Almaty, Kazakhstan; e-mail: kamshat-0707@mail.ru

Also, a similar system is studied in the field of neuroscience. In the works Huang for a class of nonlinear uncertain systems based on adaptive dynamic programming (ADP) developed methods based on neuro-observers for the optimal control problem continuous-time [9] and a neural network-based approximate optimal guaranteed cost control design [10]. Vamvoudakis et al. [11] offer a control algorithm based on ADP to solve the optimal control problem for nonlinear systems with saturating actuators and nonquadratic cost functionals. In this paper, in contrast to the works mentioned above, we consider the finite-horizon optimal control problem with fixed endpoints of trajectories. In general, the Pontryagin maximum principle determines the necessary optimality conditions and allows to find a program control, depending on the current time. In this paper we propose an approach based on sufficient conditions of optimality and Lagrange multipliers of a special type, to represent desired control in the form of synthesis control depending on the state of the nonlinear system and the current time.

This method also allows to take into account the constraints on the values of control. It should be emphasised that there is more general statement of the optimal control. The problem is considered for a three-sector economy, where the share of labour and investment resources for all three sectors of the economy can be changed simultaneously.

In practice, there are a number of optimal control problems where it is necessary to move a system from an initial state to a desired final state over a specified time interval. Such problems often arise for an economic system, when it is required to achieve a certain level of economic development on a given planning horizon.

2. Three-sector model of the economy

We consider the optimal control problem for a three-sector economy model consisting of the material sector (i = 0), the capital generating sector (i = 1) and the consumer sector (i = 2). It is assumed that each sector produces its aggregate product: the material sector produces objects of labour (fuel, electricity, raw materials and other materials); the capital-generating sector produces means of labour (machines, equipment, industrial buildings, etc.); the consumer sector produces consumer goods.

In accordance with [1], the mathematical model consists of:

- three Cobb-Douglas type functions of the product specific output:

$$x_i = \theta_i A_i k_i^{\alpha_i}, \quad A_i > 0, \ \ 0 < \alpha_i < 1, \ \ (i = 0, 1, 2). \tag{1}$$

- three differential equations describing the dynamics of the capital-labour ratio:

$$k_i = -\lambda_i k_i + \frac{s_i}{\theta_i} x_1, \quad k_i(0) = k_{i0}, \quad \lambda_i > 1, \quad (i = 0, 1, 2). \tag{2}$$

- three balance equations:

$$s_0 + s_1 + s_2 = 1, \ s_i \geq 0, \ (i = 0, 1, 2), \tag{3}$$

$$\theta_0 + \theta_1 + \theta_2 = 1, \ \theta_i \geq 0, \ (i = 0, 1, 2), \tag{4}$$

$$(1 + \beta_0) x_0 = \beta_1 x_1 + \beta_2 x_2, \ \beta_i \geq 0, \ (i = 0, 1, 2). \tag{5}$$

Here the state of the system (capital-labour ratio) is described by a vector k_0, k_1, k_2; $(s_0, s_1, s_2, \theta_0, \theta_1, \theta_2)$ is a control vector, where $s_0, s_1, s_2 -$ are shares of sectors in an investment resources distribution and $\theta_0, \theta_1, \theta_2 -$ are shares of sectors in a workforce distribution;

$x_i -$ are products specific outputs (the number of products in *i*-th sector per worker);

$\alpha_i -$ is the coefficient of elasticity of funds;

$\lambda_i -$ the capital-labour ratio;

$\beta_i -$ are direct material costs of manufacturing products in the *i*-th mentioned sector of a cluster $(i = 0, 1, 2)$.

We will consider the problem of transition of a nonlinear system from an initial state (k_{00}, k_{10}, k_{20}) to a state (k_{0s}, k_{1s}, k_{2s}) over the time interval $[0, T]$. As a desired final state (k_{0s}, k_{1s}, k_{2s}) we choose a steady state of the system, which is determined by equating the right sides of differential equations (2) to zero:

$$k_{0s} = \frac{s_0 \theta_1 A_1 k_{1s}^{\alpha_1}}{\lambda_0 \theta_0}, \quad k_{1s} = \left(\frac{s_1 A_1}{\lambda_1}\right)^{\frac{1}{1-\alpha_1}}, \quad k_{2s} = \frac{s_2 \theta_1 A_1 k_{1s}^{\alpha_1}}{\lambda_2 \theta_2}. \tag{6}$$

The values of the capital-labour ratio k_{is} $(i = 0, 1, 2)$ in the equilibrium state (6) depend on the controls $(s_0, s_1, s_2, \theta_0, \theta_1, \theta_2)$, for which we can select the values $(s_{0s}, s_{1s}, s_{2s}, \theta_{0s}, \theta_{1s}, \theta_{2s})$, the result of solving the nonlinear programming problem in order to maximise the specific consumption: $x_2 \to \max$.

2.1. Statement of the problem without limited control

The system of differential equations (2) in vector form is:

$$\dot{y}(t) = Ay(t) + Bf\big(y(t)\big)u(t), \qquad t \in [0, T], \tag{7}$$

where

$$y_1 = k_1, \quad y_2 = k_2, \quad y_3 = k_0,$$

$$u_1 = s_1, \quad u_2 = \frac{s_2\theta_1}{\theta_2}, \quad u_0 = \frac{s_0\theta_1}{\theta_0},$$

$$f(y) = f_1(y_1) = k_1^{\alpha_1}, \quad f_2(y_2) = k_2^{\alpha_2}, \quad f_3(y_3) = k_0^{\alpha_0},$$

$$A = \begin{pmatrix} -\lambda_1 & 0 & 0 \\ 0 & -\lambda_2 & 0 \\ 0 & 0 & -\lambda_0 \end{pmatrix}, \qquad B = \begin{pmatrix} A_1 & 0 & 0 \\ 0 & A_1 & 0 \\ 0 & 0 & A_1 \end{pmatrix}.$$

Here $y = (y_1, y_2, y_3)^*$ is a vector of the object state, $u = (u_1, u_2, u_3)^*$ is the control vector. The initial and final states of the system are given:

$$y(0) = y_0, \quad y(T) = y_s. \tag{8}$$

Note that the desired final state of the system $y(T) = y_s$ is an equilibrium state in which per capita consumption is maximised and ensures a balanced growth of the economy.

We consider the following problem: it is required to find a control $u(t, y)$ that moves the system (7) from a given initial state $y(0) = y_0$ to an equilibrium state $y(T) = y_s$ for over the time interval $[0, T]$, and minimising cost functional:

$$J(u) = \frac{1}{2}\int_0^T \{[y - y_s]^* Q[y - y_s] + [f(y)u - f_s u_s]^* R[f(y)u - f_s u_s]\}dt, \tag{9}$$

where Q and R are positive semidefinite and positive definite (3×3)-matrices respectively.

To solve this problem we construct an auxiliary functional with Lagrange multipliers of a special kind. For that, we add a system of differential equations (7) with multiplier $\lambda = K(t)(y(t) - y_s) + q(t)$ to the expression for functional (9). As a result, we get the following functional [12]:

$$L(y, u) = \int_0^T \{\frac{1}{2}[y - y_s]^* Q[y - y_s] + \frac{1}{2}[f(y)u - f_s u_s]^* R[f(y)u - f_s u_s] +$$

$$+ [K(t)(y - y_s) + q(t)]^* [A(y - y_s) + B(f(y)u - f_s u_s) - \dot{y}(t)]\}dt,$$

where $q(t)$ is a vector of dimension (3×1); $K(t)$ is a symmetric positive definite (3×3)-matrix.

Theorem 1. The pair of functions in problem (7)–(9) is optimal if and only if:

1. $y(t)$ satisfies the differential equation:

$$\dot{y}(t) = -A_1(t)(y(t) - y_s) + B_1 q(t), \qquad y(t_0) = y_0, \quad y(T) = y_s.$$

2. Control $u(t)$ is defined as follows:

$$u(y,t) = f^{-1}(y)\{f_s u_s - R^{-1}B^*[K(t)(y(t) - y_s) + q(t)]\}.$$

$K(t)$ is differential matrix Riccati equation.

2.2. Statement of the problem with limited control

We write the system of differential equations (7) in vector form using the following notation:

$$\dot{y}(t) = Ay(t) + BD\big(y(t)\big)u(t), \qquad y(t_0) = y_0, \quad t \in [0,T], \tag{10}$$

$$D(y(t)) = \begin{pmatrix} y_1^{\alpha_1} & 0 & 0 \\ 0 & y_1^{\alpha_1} & 0 \\ 0 & 0 & y_1^{\alpha_1} \end{pmatrix}.$$

Here, $y(t) = (y_1, y_2, y_3)^*$ is a state vector of the object, $u(t) = (u_1, u_2, u_3)^*$ is a control vector. The components of the control vector $u(t) = (u_1, u_2, u_3)^*$ satisfy two-sided constraints of the following form:

$$\gamma_1 \le u \le \gamma_2, \qquad 0 < \gamma_{1i} \le u \le \gamma_{2i} < 1, \qquad (i = 0,1,2), \tag{11}$$

which are obtained from the initial constraints (3), (4).

Murzabekov et al [13] considered optimal control problems with fixed ends of trajectories for a linearised system of an economic cluster.

It is required to find a synthesising control $u(y,t)$ that takes the system (10) from a given initial state $y(0) = y_0$ to the desired equilibrium state $y(T) = y_s$ in a time interval $[0,T]$, while minimising the functional:

$$J(u) = \frac{1}{2}\int_0^T \{[y - y_s]^* Q[y - y_s] + [D(y)u - D_s u_s]^* R[D(y)u - D_s u_s]\}dt +$$

$$+ \frac{1}{2}[y(T) - y_s]^* F[y(T) - y_s], \tag{12}$$

where Q is a positively semidefinite (3×3)- matrix, and R, F, $D(y(t))$ are positive definite matrices of dimension (3×3), $D_s = D(y_s)$. The symbol (*) means the operation of transposing a matrix or a vector.

To solve the problem, we add to the expression for the functional (12) the system of differential equations (10) with the factor $\lambda = K(t)(y(t)-y_s)+q(t)$, as well as the following expression:

$$\lambda_1^*(t)D^*(y)RD(y)[\gamma_1 - u(t)] + \lambda_2^*(t)D^*(y)RD(y)[u(t)-\gamma_1] +$$

$$+\lambda_3^*(t)[y(t) - y_s - W(t,T)q(t)],$$

where $\lambda_1(t)\geq 0,\ \lambda_2(t)\geq 0$. As a result, we obtain the following functional:

$$L(y,u) = \int_0^T \{\frac{1}{2}[y-y_s]^*Q[y-y_s] + \frac{1}{2}[D(y)u - D_s u_s]^* R[D(y)u - D_s u_s] +$$

$$+ [K(t)(y-y_s)+q(t)]^*[A(y-y_s) + B(D(y)u - D_s u_s) - \dot{y}(t)] +$$

$$+\lambda_1^*(t)D^*(y)RD(y)[\gamma_1 - u(t)] + \lambda_2^*(t)D^*(y)RD(y)[u(t)-\gamma_1] +$$

$$+ \lambda_3^*(t)[y(t) - y_s - W(t,T)q(t)]\}dt,$$

where $q(t)$ is a vector of dimension $(n\times1)$, and $K(t)$ a symmetric positive definite $(n\times n)$-matrix of dimension.

Theorem 2. Let Q be a positively semidefinite matrix, and R, F, $D(y)$ be positive definite matrices in the time interval $t_0 \leq t \leq T$; the matrix $W_0 = W(t_0,T)$ is positive definite. Suppose that system (10) is completely controllable at the instant time t_0. Then for the optimality of the pair ($y(t)$,$u(t)$) in problem (10) – (12), it is necessary and sufficient that:

1. $y(t)$ satisfies the differential equation:

$$\dot{y}(t) = A_1(t)(y(t)-y_s) - B_1 q(t) + BD(y)\varphi(y,t),\ \ y(t_0) = y_0, \quad (13)$$

2. Control $u(t)$ is defined as follows:

$$u(y,t) = D^{-1}(y)\{D_s u_s - R^{-1}B^*[K(t)(y(t)-y_s) + q(t)]\} + \varphi(y,t). \quad (14)$$

$K(t)$ is differential matrix Riccati equation.

2.3. Algorithm of solving

We describe an algorithm for solving the optimal control problem (1)–(5), that can conveniently be implemented with a computer:

Step 1. Integrate the system of differential equations $K(t)$ and $W(t,T)$ over the interval $[0,T]$ under conditions $K(T) = K_T$ and $W(T,T) = (F - K_T)^{-1}$.

Step 2. Set the conditions $y(0) = y_0$ and calculate $q_0 = W^{-1}(0,T)(y_0 - y_s)$.

Step 3. Integrate the system of differential equations (13) in interval $[0,T]$ with the initial conditions $y(0) = y_0$, $q(0) = q_0$. It is possible to output the results and a graph (if needed) of the optimal trajectory $y(t)$ and optimal control $u(t)$ in the process of integration of the system (13).

Step 4. Assume that the state of the system $y(t)$ and the optimal control $u(t)$ are found, then relations:

$$f_i(y_i) = y_i^{\alpha_i}, \qquad (i = 0, 1, 2),$$

$$v = \frac{\beta_1 A_1 f_1(y_1) + \beta_2 A_2 f_2(y_2) \dfrac{1 - u_1(t)}{u_2(t)}}{(1 - \beta_0) A_0 f_3(y_3) \dfrac{1 - u_1(t)}{u_3(t)} + \beta_2 A_2 f_2(y_2) \dfrac{1 - u_1(t)}{u_2(t)}},$$

ensure the fulfillment of conditions (5);

$$s_0 = v\big(1 - u_1(t)\big), \quad s_1 = u_1(t), \quad s_0 = (1 - v)\big(1 - u_1(t)\big),$$

ensure the fulfillment of conditions (3);

$$\theta_0 = \frac{v(1-s_1)\theta_1}{u_3(t)}, \quad \theta_1 = \frac{1}{1+\frac{s_0}{u_3(t)}+\frac{s_2}{u_2(t)}}, \quad \theta_2 = \frac{(1-v)(1-s_1)\theta_1}{u_2(t)},$$

ensure the fulfillment of conditions (4).

3. Results

Numerical calculations were computed with the values specified in Table 1.

Table 1. Parameter values for a three-sector model of the economy

i	α_i	β_i	λ_i	A_i	s_{is}	θ_{is}	k_{is}
0	0.46	0.39	0.05	6.19	0.2763	0.3944	966.4430
1	0.68	0.29	0.05	1.35	0.4476	0.2562	2410.1455
2	0.49	0.52	0.05	2.71	0.2761	0.3494	1090.1238

We also solve the problem of optimal control for the values of the initial state of the system $y(t_0)$, which were chosen in the following form:

$$y(t_0) = (-600, -200,\ 200)^*, \tag{15}$$

Matrices Q and R were chosen as:

$$Q = \begin{pmatrix} 0.0277 * 10^{-4} & 0 & 0 \\ 0 & 0.25 * 10^{-4} & 0 \\ 0 & 0 & 0.25 * 10^{-4} \end{pmatrix},$$

$$R = \begin{pmatrix} \frac{1}{8100} & 0 & 0 \\ 0 & \frac{1}{1600} & 0 \\ 0 & 0 & \frac{1}{1250} \end{pmatrix},$$

Matrix K_T is equal to:

$$K_T = \begin{pmatrix} 0.107 * 10^{-4} & 0 & 0 \\ 0 & 0.77 * 10^{-4} & 0 \\ 0 & 0 & 0.85 * 10^{-4} \end{pmatrix}.$$

The results of the system state calculations are shown in Figures 1 and 2. It can be seen from Figure 3 that the optimal controls do not go beyond the range of the U determined by the constraints (11). For the example under consideration, these restrictions have the form:

$$0.1 \le u_i \le 0.9, \qquad (i = 0, 1, 2). \tag{16}$$

In Figure 3 the control components $u_1(t)$ and $u_3(t)$ lie on the boundary of the region U in the time interval $[0, t_1]$, $[0, t_2]$, respectively, then for $t_1 \in (t_1, T]$, $t_2 \in (t_2, T]$, the lines enter the interior of the region U Switching controls occurs at times $t_1 = 0.637$, $t_2 = 2.234$ for components $u_1(t)$, $u_3(t)$ respectively. The time interval $[t_0, T] = [0, 20]$.

Figures 4 and 5 show the changes in resources that ensure the system (10) is brought to an equilibrium state and satisfies the balance ratios (3)–(5).

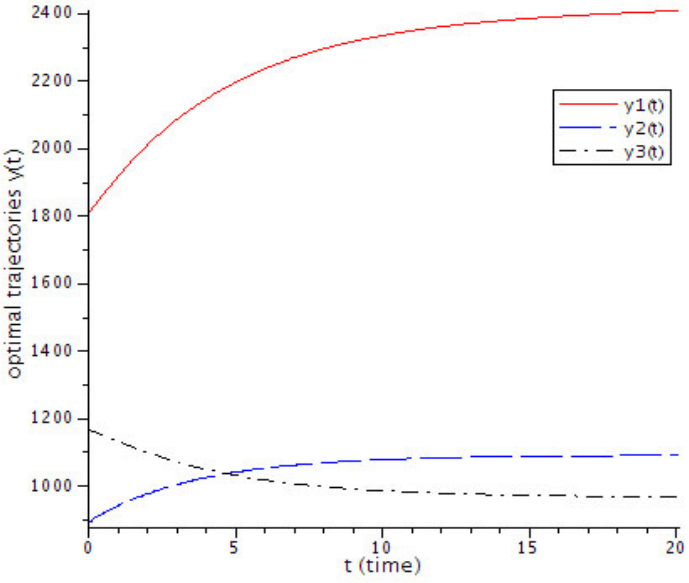


Fig. 1. Graphs of optimal trajectories for the system (10) with the initial condition (15) under the control (14)

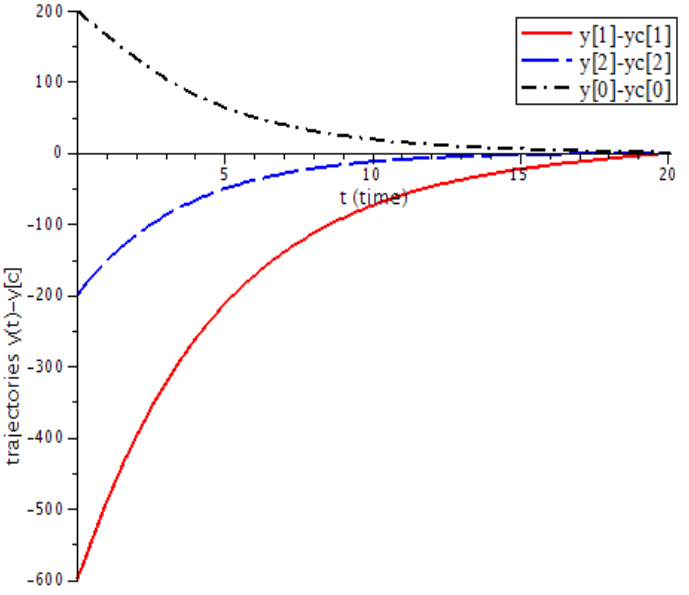


Fig. 2. Graphs of trajectories $y(t) - y_s$

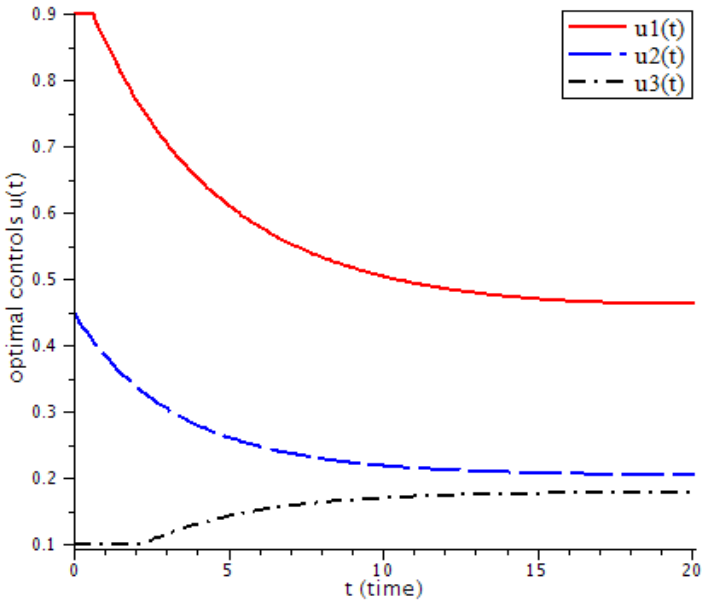


Fig. 3. Graphs of the optimal controls for the system (10) with the initial condition (15)

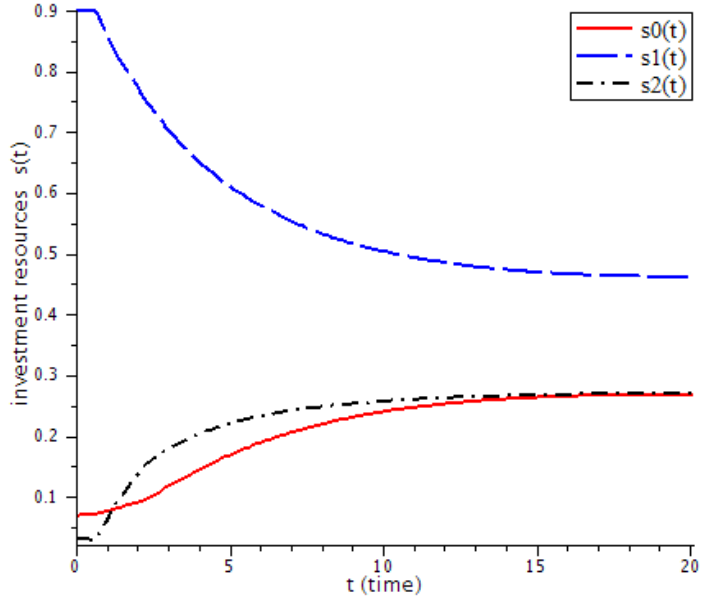


Fig. 4. Graphs of the optimal allocations of investment resources for the balance ratios (3)–(5) for the system (10) with the initial condition (15)

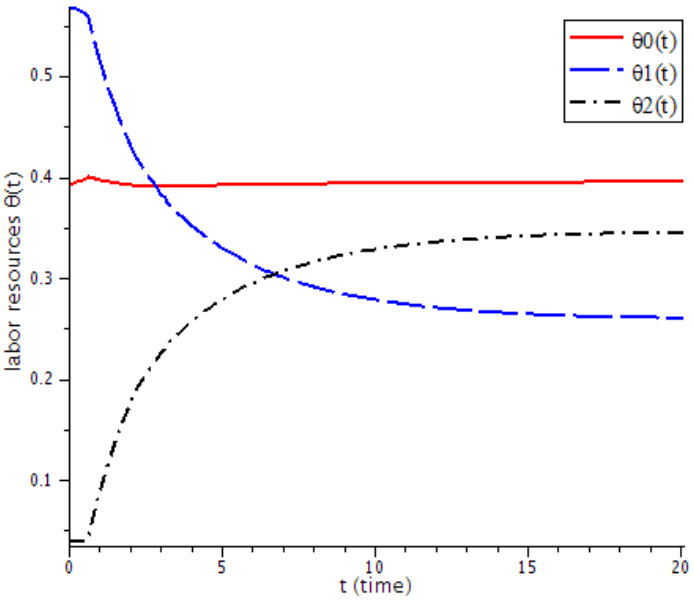


Fig. 5. Graphs of the optimal distribution of labour resources for the balance ratios (3)–(5) for the system (10) with the initial condition (15)

4. Conclusions

In this paper we propose a new approach of constructing a synthesis control, based on the principle of feedback with constraints on the control and system status, that moves a dynamic system to a desired state in a finite time.

The problem is solved using Lagrange multipliers depending on the phases' coordinates and time. By choosing the proper multiplier it is possible to construct the optimal control based on the feedback principle, followed by selection of the control parameters to satisfy algebraic conditions on the control and system status.

Numerical examples of the initial condition are considered. Figures 1–5 shows the optimal trajectories when with the chosen initial condition in the form (15). In this case, the controls $u_1(t)$, $u_2(t)$, $u_3(t)$ take values satisfying the given constraints (14) and the balance relations (3)–(5) are satisfied.

Bibliography

[1] Kolemayev V.A., Optimal balanced growth of open three-sector economy (in Russian), Prikladnaya Econometrika, 2008, 11(3): 15–42.

[2] Zhang J.S., The analytical solution of balanced growth of non-linear dynamic multi-sector economic model, Economic modeling, 2011, 28(1): 410–421.

[3] Aseev S.M., Besov K.O., Kryazhimskii A.V., Infinite-horizon optimal control problems in economics, Russian Mathematical Surveys, 2012, 67(2): 195–253.
[4] Klamka J., Constrained controllability of dynamics systems, International Journal of Applied Mathematics and Computer Science, 1999, 9(2): 231–244.
[5] Klamka J., Controllability of nonlinear discrete systems, International Journal of Applied Mathematics and Computer Science, 2002, 12(2): 173–180.
[6] Miłosz M., Murzabekov Z., Tussupova K., Usubalieva S., Optimisation of Discrete Processes with Bounded Control, Journal of Information Technology and Control, 2018, 47(4): 684–690.
[7] Mitkowski W., Bauer W., Zagórowska M., Discrete-time feedback stabilization, Archives of Control Sciences, 2017, 27(2): 309–322.
[8] Afanas'ev V.N., Orlov P.V., Suboptimal control of feedback-linearizable nonlinear plant, Journal of Computer and Systems Sciences International, 2011, 50(3): 365–374.
[9] Huang Y., Neuro-observer based online finite-horizon optimal control for uncertain non-linear continuous-time systems, IET Control Theory & Applications, 2017, 11(3): 401–410.
[10] Huang Y., Optimal guaranteed cost control of uncertain non-linear systems using adaptive dynamic programming with concurrent learning, IET Control Theory & Applications, 2018, 12(8): 1025–1035.
[11] Vamvoudakis K.G., Miranda M.F., Hespanha J.P., Asymptotically stable adaptive-optimal control algorithm with saturating actuators and relaxed persistence of excitation, IEEE Transactions on Neural Networks and Learning Systems, 2016, 27(11): 2386–2398.
[12] Aipanov Sh.A., Murzabekov Z.N., Analytical solution of a linear quadratic optimal control problem with control value constraints, Journal of Computer and Systems Sciences International, 2014, 53(1): 84–91.
[13] Murzabekov Z., Milosz M., Tussupova K., The Optimal Control Problem with Fixed-End Trajectories for a Three-Sector Economic Model of a Cluster, ACIIDS 2018, 382–391. Springer Verlag: Heidelberg, Germany.

Dauren Nazarbayev[1], Zhanna Alimzhanova[2]

Adaptation of dynamic models to cryptanalysis conditions

Abstract: The report provides an overview of dynamic models that adapted to the conditions of cryptanalysis, make it possible to obtain those characteristics that reflect the peculiarities of encryption. The basic principles of internal interconnection of cryptanalysis and some aspects of the theory of dynamical systems are described. One of the main characteristics of dynamic models adaptable to cryptanalysis is to identify the frequency of the encoded information. A cryptographic encryption system in dynamic systems is considered. Dynamic models are currently used in many areas of scientific research, where time parameters are used that connect the past, present and future of processes. We can predict the behaviour of processes or phenomena using data from dynamic models. In this scientific work, the choice of dynamic models for cryptanalysis conditions is due to the fact that models with time parameters have a rich historical, research scientific base, this theory of dynamical systems is a separate scientific branch of applied mathematics, on which outstanding mathematicians of the last century worked. Dynamical systems are also related to other areas of mathematics, with algebra, topology, logic, probability theory, geometry, number theory, etc. If we continue talking about dynamical systems in general, then this theory is applied, as we know, in many areas of scientific activity in physics, economics, engineering, technology, sociology, computer science, and it began to be used in information security.

Keywords: dynamical systems, information security, pseudo-random system generators, cryptography, cryptanalysis, cyclicality, Fourier transform, Kotelnikov theorem

1. Introduction

The theory of dynamical systems deals with the evolution of systems. It describes processes in motion, tries to predict the future of these systems or processes and understand the limitations of these predictions [1].

Dynamical systems have three components: phase space (state space), time and evolution law. Phase space is a space in which all its components have a possible state of the system at a certain period or moment in time. If we talk about time, then time is divided into discrete and continuous, and their set of values, respectively, are whole numbers and real numbers. The law of evolution is a law in which one can assess behaviour at any moment in time.

The most important advantage of choosing dynamic systems is that all systems in this category have some important properties that make them widely applicable. One of these important properties is its stability of processes against

[1] Dauren Nazarbayev, Al-Farabi Kazakh National University, 71 al-Farabi Ave., Almaty, Kazakhstan; e-mail: d.a.nazarbayev@gmail.com

[2] Zhanna Alimzhanova, Al-Farabi Kazakh National University, 71 al-Farabi Ave., Almaty, Kazakhstan; e-mail: zhannamen@mail.ru

various external disturbances. By sustainability, we mean the constancy of the foundation of the structure and the basic functions performed over a period of time, with some comparisons for the invariability of environmental conditions. Let us give examples of dynamical systems from the field of mathematics, maybe a logistic mapping, the Collatz problem, cellular automata, billiards, a geodesic flow, and many other systems [1].

Examples of dynamic systems in information security are dynamic systems of chaotic behaviour, which are used in many other fields of science. These systems are based on the idea of creating pseudo-random number generators. Pseudo-random system generators are used in cryptography to generate secret keys that play a fundamental role in the encryption of some plain texts.

Table 1. Classes of dynamical systems

Time G (semigroup)	**Action**
Natural numbers (N,+)	Maps
Integers (Z,+)	Invertible maps
Positive real numbers (R+,+)	Semiflows (some PDE's)
Real numbers (R,+)	Flows (Differential equations)
Any group (G, ⋆)	Representations
Lattice (Zn,+)	Lattice gases, Spin systems
Euclidean space (Rn,+)	Tiling dynamical systems
Free group (Fn, ◦)	Iterated function systems

Source: O. Knill [1]

We mention the general definition to stress that the ideas developed for one dimensional time generalize to other situations. Because physical time is one dimensional, the important cases for us are definitely discrete and continuous dynamical systems:

- dynamics of maps defined by transformations;
- dynamics of flows defined by differential equations.

All areas of mathematics are linked together in some way or an other. Intersections of fields like algebraic topology, geometric measure theory, geometry of numbers or algebraic number theory can be considered full blown independent subjects. The theory of dynamical systems has relations with all other main fields and intersections typically form subfields of both.

- Algebra
- Topology
- Logic
- Measure theory
- Probability theory
- Dynamics
- Analysis
- Geometry
- Number Theory

Examples of dynamical systems:

- the logistic map;
- the Lorentz system;
- the Collatz problem;
- computing square roots;
- cellular automata;
- differential equations in the plane (e.g. Van Der Pool Oscillator);
- billiards;
- standard map;
- geodesic flow;
- the henon map;
- solving equations;
- three body problem;
- exterior billiards;
- the digits of PI;
- lattice points near graphs.

In the field of information security, all applicable systems that participate in encryption must also be reversible systems, which are one of the essential conditions for encryption. Many dynamical systems are constructed in the form of some differential equations that can be approximated in discrete difference schemes, since any architecture of technical devices has some finite bit depth. Thus, we can say that any input and output data are limited to some sequences of bits, any finite difference scheme can be represented as some finite state machine [2].

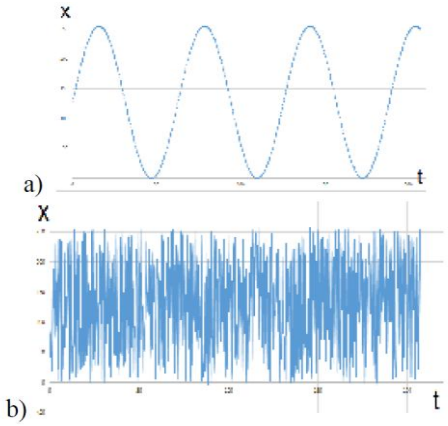


Fig. 1. Periodic signal encryption (a – original, b – encrypted, signal transmission time, x – amplitude)

Source: V.V. Kirichenko, E.V. Lesina [2]

Many dynamic systems with I/O traffic are signal transformations. A certain discrete message is the input data, and the encrypted data sequences transmitted over telecommunication channels are the output data. The encryption of the sequence in the form of text follows the following scheme:

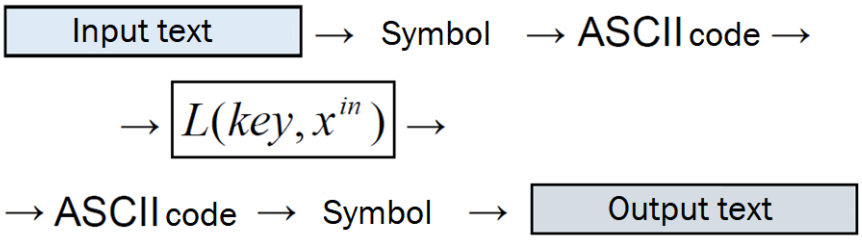


Fig. 2. The encryption of the sequence in the form of text follows the following scheme

Source: V.V. Kirichenko, E.V. Lesina [2]

The code is taken that corresponds to a certain character from the ASCII code, which in turn changes into another character by the automaton $L(key, x^{in})$ [2].

On the other hand, any dynamic system that has an I / O structure can be used primarily for signal transformation. As a result, the character becomes a text character thanks to the character table. So, from the text sequence we transform it into another encrypted sequence representing the signal. The other way round, the decryption of the signal is carried out according to the rules according to which it was carried out by an automatic machine, but now in this case already by a reverse automatic machine. In this case, it is still unknown to talk about the effectiveness of this algorithm, and this is the main point. The idea of reverse systems in the information security, for example, access control, are the tasks of the synthesis of modern algorithms [3].

Transmission channels often do not have anti-interference protection for the average user and it is very convenient for a single module to perform many functions, for this task it is necessary to use cyclic codes [4].

2. Identification of cyclicality through time characteristics

As we know, any processes in nature are characterized by their continuity or discreteness. But, it is convenient for us to round or divide continuous processes or signals into finite values, and we cannot work with continuous ones, and therefore we have to refer to discrete sequences by transforming them from continuous sequences of signals or information.

Thus, we have learned to discretize any processes or phenomena into some information, for example, if we talk about time, then time is reflected in some watch mechanisms. Any information exchange is limited to a certain symbolic sequence. So, we have learned how to move from continuous sequences of any values to discrete ones. But the converse is also important, the transition from discrete sequences to continuous ones is correct, says the Kotelnikov theorem. Suppose some continuous process is represented in the form of some vibration of various kinds of frequencies. To restore the process of oscillations of the same frequency, it is required to find the component part of the oscillations of the highest frequency. Thus, from discrete points, we can restore the original process of the same frequency. If the analog signal in the system is sampled uniformly at a frequency that exceeds the highest signal frequency by at least two times, then the outgoing analog signal will be completely restored from the discrete sequence of values.

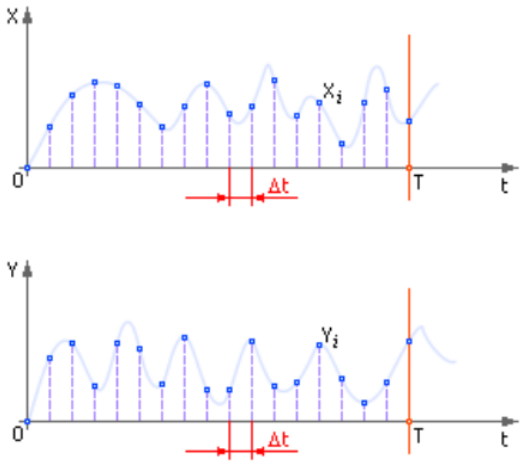


Fig. 3. Input signal X versus time t and output signal Y versus time t

Source: O.I. Mukhin [9]

To evaluate the dynamic performance of the information protection system against unauthorized access in automated information systems, there is an indicator of “time efficiency” that describes the security of automated information systems as a whole and should be understood as “the ability of the information protection system against unauthorized access to carry out a given action in a time interval that meets the specified requirements” [5].

3. Example of creating a dynamic model of information security threats

How can the dynamic model be applied in data protection? For example, consider such a concept as the confidentiality of information. Confidentiality of information – the property of hiding information from outsiders, who should not know the content of the information. If an outsider learns the content of the information, it is considered that the confidentiality of the information is violated. To build a model of information security threats, it is necessary to create some designations in advance, for example, let us have I – this is the amount of information in the system, dI – the flow of information outside the information system, $\frac{dI}{dt}$ – the rate of change of this flow. If

$$dI = 0; \frac{dI}{dt} = 0, \tag{1}$$

then this expression is interpreted to mean that there is no information leak. Information leakage depends on the security of the system. Let's make the following expression:

$$\frac{dI}{dt} = C_{d1} * D - C_v * I, \tag{2}$$

where D is an indicator of information security, C_v – a coefficient that reflects the impact of the amount of information on its leakage; C_{d1} – a coefficient that reflects the impact of security on information leakage. Thus, from the equation we see that the leak is dependent on the size of the information system and security on information depends on the size of the system and security threats information.

$$\frac{dD}{dt} = D_p - C_{d2} * I - V_d * I. \tag{3}$$

Combining the equations, we obtain a dynamic model of system security threats.

$$\begin{cases} \frac{dI}{dt} = C_{d1} * D - C_v * I \\ \frac{dD}{dt} = D_p - C_{d2} * I - V_d * I \end{cases}. \tag{4}$$

Let's find the stationary position of the system described by equations (4). Stationarity conditions $dI = 0; \frac{dI}{dt} = 0$. Based on this:

$$\begin{cases} C_{d1} * \bar{D} - C_v * \bar{I} = 0 \\ D_p - C_{d2} * \bar{I} - V_d * I = 0. \end{cases} \quad (5)$$

From the second equation of the system:

$$\bar{I} = \frac{D_p}{C_{d2}+V_d}. \quad (6)$$

Next, let's derive $\bar{D}$ from the first equation of the system of equations (5).

$$C_{d1} * \bar{D} - \frac{C_v * D_p}{C_{d2}+V_d} = 0 \ , \quad (7)$$

$$\bar{D} = \frac{C_v * D_p}{C_{d2}+V_d} * \frac{1}{C_{d1}}. \quad (8)$$

From here, let's derive the following conditions for the stationarity position of the system:

$$\begin{cases} \bar{I} = \frac{D_p}{C_{d2}+V_d} \\ \bar{D} = \frac{C_v * D_p}{C_{d2}+V} * \frac{1}{C_{d1}} \end{cases}. \quad (9)$$

Let's consider a variant of solving the system of equations by the method of "small deviations" $I = \bar{I} + I; D = \bar{D} + D$ which gives followings:

$$\begin{cases} \frac{dI}{dt} = C_{d1} * (\bar{D} + D) - C_v * (\bar{I} + I) \\ \frac{dD}{dt} = D_p - C_{d2}(\bar{I} + I) - V_d(\bar{I} + I) \end{cases}, \quad (10)$$

$$\begin{cases} \frac{dI}{dt} = C_{d1} * \left(\frac{C_v * D_p}{C_{d2}+V_d} * \frac{1}{C_{d1}} + D\right) - C_v * (\frac{D_p}{C_{d2}+V_d} + I) \\ \frac{dD}{dt} = D_p - C_{d2}\left(\frac{D_p}{C_{d2}+V_d} + I\right) - V_d(\frac{D_p}{C_{d2}+V_d} + I) \end{cases}, \quad (11)$$

$$\begin{cases} \frac{dI}{dt} = \frac{C_{d1} * C_v * D_p}{C_{d2}+V_d} * \frac{1}{C_{d1}} + C_{d1} * D - \frac{C_v * D_p}{C_{d2}+V_d} - C_v * I) \\ \frac{dD}{dt} = D_p - \frac{C_{d2} * D_p}{C_{d2}+V_d} - C_{d2} * I - \frac{V_d * D_p}{C_{d2}+V_d} - V_d * I \end{cases}, \quad (12)$$

$$\begin{cases} \frac{dI}{dt} = C_{d1} * D - C_v * I \\ \frac{dD}{dt} = -I(C_{d2} + V_d) \end{cases}, \quad (13)$$

Let's differentiate the first equation:

$$\frac{d^2I}{dt^2} = -I * C_{d1} * (C_{d2} + V_d) - C_v * \frac{dI}{dt}, \tag{14}$$

$$\frac{d^2I}{dt^2} + C_v * \frac{dI}{dt} + I * C_{d1} * (C_{d2} + V_d) * I = 0. \tag{15}$$

This (15) is the equation of a harmonic oscillator with a damped amplitude.

Let's analyze the behavior of the system by converting from the differential form of equations (4), (5) to the discrete form and simulate a certain interval of the system's existence. Like this:

$$\begin{cases} \frac{I_{n+1}-I_n}{\Delta t} = C_{d1} * D_n - C_v * I_n \\ \frac{D_{n+1}-D_n}{\Delta t} = D_p - C_{d2} * I_n - V_d * I_n \end{cases}, \tag{16}$$

$$\begin{cases} I_{n+1} = I_n + (C_{d1} * D_n - C_v * I_n) * \Delta t \\ D_{n+1} = D_n + \left(D_p - I_n * (C_{d2} + V_d)\right) * \Delta t \end{cases}. \tag{17}$$

We can set initial parameters and visualize the results using this equation. Based on the results, we can make an assessment of the system's behavior under different initial conditions and information security parameters.

A modeling method is considered, on the basis of which it is possible to assess the stability of the information security system [6].

Further research is planned to model and analyze more specific dynamic systems for cryptanalysis.

Given that this model is abstract and simple, it is necessary to study new parameters that can be added to a model that approximates real systems. The relevance of the use of dynamic models is the starting point for effective measures to protect information in the system as a whole [7].

Threat classification models are analyzed and arguments are presented about the lack of completeness and validity of existing threat classification models. It proposes a scalable approach that classifies security threats and a mixed threat classification model that allows the definition and articulation of the characteristics of the threat [8].

4. Fourier transform in dynamical systems

This modeling method is characterized by the fact that harmonics exist in all signals. When summing these harmonics, it is a signal model. Suppose that a certain signal contains the sum of three harmonics: $4\cos(t) + 3\cos(2t) + \cos(3t)$.

Fig. 4. Example of a harmonic signal

Source: O.I. Mukhin [9]

A complex signal can be displayed as a sum of harmonics. Now let's analyze a signal that is periodic, then its time parameter is equal to a period of some *p*, otherwise, if the signal is non-periodic, then the entire signal interval can be considered a period here [9].

$$H_0 = \frac{2}{p}\int_0^p X(t)dt,$$

$$H_1 = \frac{2}{p}\int_0^p X(t)\cos\left(\frac{2\pi t}{p}\right)dt, \quad G_1 = \frac{2}{p}\int_0^p X(t)\sin\left(\frac{2\pi t}{p}\right)dt,$$

$$H = \frac{2}{p}\int_0^p X(t)\cos\left(\frac{2\pi 2t}{p}\right)dt, \quad G_2 = \frac{2}{p}\int_0^p X(t)\sin\left(\frac{2\pi 2t}{p}\right)dt, \qquad (18)$$

…

$$H_i = \frac{2}{p}\int_0^p X(t)\cos\left(\frac{2\pi it}{p}\right)dt, \quad G_i = \frac{2}{p}\int_0^p X(t)\sin\left(\frac{2\pi it}{p}\right)dt.$$

H_i, G_i – the coefficients of harmonics, called weights, i – numbering of harmonics. (18) – direct Fourier transform. $\frac{2\pi \mathrm{i}}{p} = \omega_i$ – harmonic frequency.

Note that the following $\omega_\mathrm{i} = \mathrm{i}\omega_1$ dependence is observed.

To restore the original signal, the inverse Fourier transform formula is used, in which weights $(H_0, H_1, H_2, \ldots, G_1, G_2, \ldots)$ are present, which can be stored as a vector instead of the entire signal, are called full signal characteristics.

$$x(t) = \frac{H_0}{2} + H_1 \cos\left(\frac{2\pi t}{p}\right) + G_1 \sin\left(\frac{2\pi t}{p}\right) + H_2 \cos\left(\frac{2\pi 2t}{p}\right) +$$

$$+ G_2 \sin\left(\frac{2\pi 2t}{p}\right) + \cdots + H_i \cos\left(\frac{2\pi it}{p}\right) + G_i \sin\left(\frac{2\pi it}{p}\right) + \cdots . \qquad (19)$$

Here (19) is the inverse Fourier transform.

I would like to note the fact that these weights are used in signal processing in dynamic systems.

Suppose such that, changing the dynamic signal X(t) through object P gives the output signal Y(t). It may be that the object can be represented by some mathematical model (for example, differential equations), the input signal will be expressed by some integration, then the operation will be laborious in computer systems [10].

Suppose that we have a five-section electronic system, let all sections be linear, except for the 3-section, and for it the amplitude-frequency response is undefined.

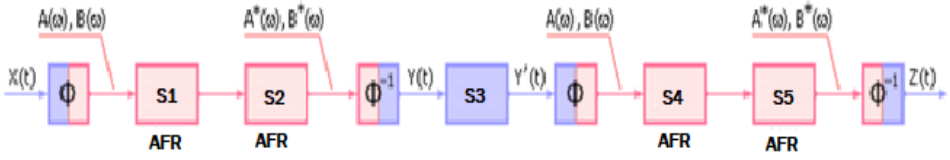


Fig. 5. Five-section electronic system

Source: O.I. Mukhin [9]

Here AFR is the amplitude-frequency characteristic.

As shown in the picture, the input signal X(t) is initially in the time domain, then the direct Fourier transform passes through the frequency domain $A(\omega), B(\omega)$, after the 2-section $A^*(\omega), B^*(\omega)$, using the inverse Fourier transform, the signal enters the time domain Y(t), passing through the 3-nonlinear section $Y^*(t)$ using the direct Fourier transform, it goes back into the frequency domain $A(\omega), B(\omega)$ passing the 5-section $A^*(\omega), B^*(\omega)$ by the inverse Fourier transform, the signal z(t) goes into the time domain, which is the result of the simulation [9].

The considered method is the fastest method. The performance lies in the replacement of some operations.

In accordance with the Kotelnikov theorem, in order not to miss a definitely corresponding harmonic, it will be necessary to sample the signal with a frequency not less than 2 times higher than the highest frequency presented in the analog signal. In other words, if the system uniformly samples an analog

signal at a frequency that is at least twice the highest frequency of the signal, the original analog signal can be completely recovered from the sampled values.

Also, if the analog signal has a finite spectrum (limited in width), then it can be restored unambiguously and without loss in discrete samples taken at a frequency strictly twice the upper frequency.

5. Conclusion

The projected dynamic model is designed to create conditions for assessing the stability of cryptosystems in order to determine vulnerabilities and the adequacy of measures to ensure an adequate level of information security. Such assessments are used in cryptanalysis of the general state of security of cryptosystems in order to make management decisions in the process of developing technical solutions and choosing crypto algorithms taking into account risk assessment.

The study of tasks for organizing information security is problematic due to the dependence on the conditions for the functioning of information systems, which are uncertain and unpredictable in various situations, and very likely in extreme situations. At the same time, previously used mathematical models give ambiguous results to assess the state and control of information protection systems operating under conditions of uncertainty. In this connection, we need new methods of modeling.

In this scientific work, it is planned to consider in the form of a base dynamic systems for modeling such systems. Using a mathematical dynamic model, as well as methods of functional analysis and representation theory, it is planned to develop a model that reveals the cyclicity of the ciphertext.

Bibliography

[1] Knill O., Dynamical systems, 2005.

[2] Kirichenko V.V., Lesina E.V., Features of using dynamic systems in data protection algorithms, problems of informatization and management, scientific journals of the National Aviation University, 2017, 3(59).

[3] Gulyaev Yu.V., Belyaev R.V., Vorontsov G.M., Information technologies based on dynamic chaos for the transmission, processing, storage and protection of information, 2000, http://www.cplire.ru.

[4] Petrenko V., Pavlov A., Ryabtsev S., Apurin A., Development of an Encryption Method Based on Cyclic Codes, 21st International Scientific Workshop on Computer Science and Information Technologies (CSIT 2019), 2019.

[5] Drovnikova I.G., Meshcheryakova T.V., Popov A.D., Rogozin E.A., Sitnik S.M., Mathematical model for estimating the efficiency of information security systems by means of Laplace transformation and Givens method, 2017, DOI 10.15622/sp.52.11, 244 (in Russian).
[6] Tutubalin P.I., Kirpichnikov A.P. Model of analysis of sustainable management of information security of a distributed information system, Vestnik tekhnologicheskogo universiteta, 2017, 20(9).
[7] Gorbylev A.L., Gorbyleva E.L., Linear dynamic model of information security threats, 2018, 25(3).
[8] Novohrestov A.K., Konev A.A., Shelupanov A.A., Egoshin N.S., Information and information carrier security threat model, Vestnik Irkutskogo gosudarstvennogo tekhnicheskogo universiteta, 2017, 21(10): 93–104. DOI: 10.21285/1814-3520-2017-12-93-104 (in Russian).
[9] Mukhin O.I., Modeling of Systems (electronic textbook), 2018.
[10] Agureev K.I., The application of deterministic chaos to transmission of information, Izvestiya TulGU, Tekhnicheskiye nauki, 2014, 11(2).

Diana Rakhimova [1], Aliya Turganbayeva[2], Akbota Kulzhanova [3], Alima Suleimenova[4]

Semantic analysis of the Kazakh language based on machine learning

Abstract: This work provides an overview of existing modern methods and software approaches for semantic analysis. Based on the research done, it has been revealed that, for the semantic analysis of text resources, an approach based on machine learning is most used. This article presents the developed algorithm for the semantic analysis of the text in the Kazakh language, taking into account the keywords and describes the work of the modules. When developing an algorithm to match certain information to a certain attribute, the choice was made on a neural network (NN) with a hidden layer (100). To begin with, preprocessing of the text is performed. For text preprocessing, we used the developed natural language processing modules. After applying these modules, the signs of our descriptor were extracted. A feature vector was then constructed using the extracted data. The constructed feature vector was compared with certain keywords, determined by the modified TF-IDF method for the Kazakh language. At the second stage, the neural network was trained. The paper also presents a software solution to this approach implemented in the Python programming language. The paper presents the results of the experiments in the graphical form of a set of words. The novelty of the proposed approach lies in the identification of semantic close words in meaning in texts in the Kazakh language. This work contributes to solving problems in machine translation systems, information retrieval, as well as analysis and processing systems in the Kazakh language.

Keywords: semantics analysis, machine learning, Kazakh language, NLP

1. Introduction

Semantic analysis of the text is the selection of semantic relations, the formation of semantic representation. In the general case, a semantic representation is a graph, a semantic network that reflects binary relations between two nodes – semantic units of the text.

Computer semantic analysis is closely related to the problem of text understanding by a machine. There are many interpretations of the concept "meaning of the text" and the tasks of understanding it. For example, according to D.A. Pospelov [1], the system understands the text entered into it if from the

[1] Diana Rakhimova, Al-Farabi Kazakh National University, Almaty, Kazakhstan; e-mail: di.diva@mail.ru
[2] Aliya Turganbayeva, Al-Farabi Kazakh National University, Almaty, Kazakhstan; e-mail: turganbaeva.aliya@bk.ru
[3] Akbota Kulzhanova, Al-Farabi Kazakh National University, Almaty, Kazakhstan; e-mail: akbota.kulzhanova1594@gmail.com
[4] Alima Suleimenova, Al-Farabi Kazakh National University, Almaty, Kazakhstan; e-mail: suleimenovaa@gmail.com

point of view of a person (or a group of experts) it correctly answers questions related to the information contained in the text. Here we can talk not about simply obtaining facts that are clearly present in the text, but about revealing the hidden meanings that the author introduces. D.A. Pospelov identifies several levels of text comprehension, from the point of view of the complexity of the questions that the intellectual system should be able to answer. Guided by the definition from [1], the meaning of the text can be considered as a description of the knowledge contained in it, in the formal language of knowledge representation, which allows solving a fairly wide range of problems related to text analysis, and the problem of semantic analysis – as a translation of a natural language – expressions into the language of knowledge representation. For example, the language of first-order predicates, semantic networks, frames, as well as ontologies and thesauri can act as a language for representing knowledge of a text in a natural language.

In the 60s and 70s, the main approach to representing the semantics of a language was the component approach, within which the meaning of each word in a natural language had to be represented as a combination of semantic universals. By the mid-1980s, it became clear that a generally accepted set of such universals had never been compiled. Relational semantics has become an alternative to the component approach in semantics. In this approach, the meanings of the words of the language are described by setting connections with the meanings of other words, and the entire conceptual system of the language is represented as a semantic network [2].

2. Review of methods and software approaches for semantic analysis

Of course, no software can replace the analysis that humans can think of. However, the programs that are currently being developed can reduce the time spent studying large databases. In this regard, the work of the following programs for solving problems of semantic text analysis is considered. Software offered by various manufacturers, such as *Semantic LLC*, *Tomita-parser* (*Yandex*), Semantic Analyst *JHON, SummarizeBot API, TextAnalyst 2.0, Galaktika-ZOOM, NLP ISA, Natasha* and *etc.* is used in various subject areas and for different languages [3–10]. A complete overview of existing modern systems of semantic analysis and their description are presented in Table 1 (look at the p. 50).

Table 1. Review of modern software systems for semantic text analysis

System name	**Description**
Semantic LLC	is a program for editing unstructured text. The semiconductor line is graphically oriented, each node is a semantic element, and the walls represent the elements of the elements. Each node attribute has a great value, the set of attributes depends on the element type
Tomita-parser (Yandex)	a program that allows you to extract facts from the structured text. The separation of facts is based on context-independent grammar rules. And the program requires a dictionary of keywords. The parser will write its own grammar
JHON	the semantic analyst *JHON* receives the meanings of a natural language in Russian and solves the following tasks: lexical analysis, morphological analysis, syntactic analysis, semantic analysis – involving the triad of subject-object relations, creating a semantic network of text, a fact of events
SummarizeBot API	the web service offers a RESTful API to handle all text and image processing tasks. It uses over 100 languages including Russian, English, Chinese, Japanese, and uses machine learning technology. The current version uses the following parameters: 1. automatically link to text; 2. Selection of keywords and conceptual documents; 3. Analysis of a sample of documents and selection of material objects and attributes; 4. Automatically detect the language of the document; 5. Obtaining unpublished data: the main text of articles, forums, forums, etc.; 6. Image processing: identification and recognition of objects in images
TextAnalyst 2.0	the program was developed by the research and product innovation center MicroSystems as a tool for text analysis. Text links allow you to create a semantic web of comments, expressed in the processed text. The request has the ability to semantic search for text fragments, taking into account the semantic links hidden in the text. Allows you to parse text by composing a hierarchical tree / heading topics containing text
Galaktika-ZOOM	an automated information search and analysis system manufactured by the Galaktika Corporation. It is a powerful editing and processing tool that allows you to get the information you need in large quantities. It is offered as a commercial system with consumers in advertising, government, and media. This program allows you to build semantic networks, but its program codes are not shared with the system
NLP ISA	for the text, a tree of analyzed analysis was built, the semantic role and connections were established. Allows you to select serialized syntax and semantic analysis mode. Alternatively, you can also select a mode that has a syntax-semantic mode combination
Natasha	it is a set of rules for getting a Tomita parser for Python and a set of ready-to-execute rules, addresses, terms, sums, and other objects

Scientific works [11–13] describe the basic ideas of information retrieval. Various options for finding text statistics are presented, which include counting the number of occurrences of words in documents and the frequency of word contiguity, and new model architectures for computing continuous vector representations of words from very large datasets. The quality of vector representations of words obtained by various models was studied using a set of syntactic and semantic language problems. In [14], the application of language models of a neural network to the problem of calculating semantic similarity for the Russian language is shown. The tools and bodies used, and the results achieved are described.

The above software products are designed for multi-resource languages such as English, Spanish. Russian, etc. Unfortunately, for the Kazakh language now there is no software implementation in the open access. This is since the Kazakh language differs in its semantic and linguistic properties from others, and also does not have large linguistic resources for conducting applied research.

3. Algorithm for semantic analysis of text in the Kazakh language

During digital technologies, given the constant growth of the volume of digital data, an important role is played by improving the quality of information retrieval using new semantic approaches and methods.

To work with big data, various algorithms and methods are being developed for the machine solution of this problem, since the amount of data does not allow for manual analysis. Any natural-language is complex, unique, and multifaceted in its own way, therefore, extracting data from documents and text resources is a large and time-consuming work that requires preliminary processing.

Based on the research done from the developed models used most for the semantic analysis of text resources, there is an approach based on machine learning. Below will be presented the developed algorithm for semantic analysis of text in the Kazakh language and implementation based on this approach. When developing an algorithm to map certain information to a certain attribute, we opted for a neural network (NN) with a hidden layer (100). Neural network training consists of the following parts:

1. Text preprocessing. Text preprocessing consists of three stages: tokenization, removal of stop words, normalization of words.
2. Construction of the feature vector. The feature vector is a sign of the characteristic we are interested in. For one descriptor, the features were taken as follows: a window of two words after, five before was taken in the text of the article at the place of occurrence of the element. Moreover, a dictionary is formed for each descriptor, which is responsible for the presence of the specified word in the dictionary. All features of each descriptor are collected into one and a feature vector is constructed.

3. Training the neural network. The network is trained by presenting each input dataset and then propagating the error.

At the second stage, the neural network was trained. For text preprocessing, the developed natural language processing modules were used. After applying these modules, we extracted the features of our descriptor. A feature vector was then constructed using the extracted data. The constructed feature vector was compared with certain keywords, determined by the modified TF-IDF method for the Kazakh language.

4. Software solution and algorithm implementation

This is one of the most difficult and demanded tasks facing artificial intelligence is NLP (Natural Language Processing). To solve and implement NLP tasks currently, there are several software systems and libraries, which include the tasks of speech recognition, language formation, and information acquisition, etc.

Python is currently one of the most promising programs for solving NLP problems. Libraries written in Python are designed to solve NLP problems and allow you to simulate various languages and processing functions.

There are also many types of libraries, consider the most famous and applicable for word processing tasks: Spacy, NLTK, CoreNLP, StanfordNER, etc. Table 2 below shows a comparison of the functional capabilities for solving the NLP problem.

Table 2. Comparison of the capabilities of libraries aimed at solving NLP problems

Function	**Spacy**	**NLTK**	**CoreNLP**
Programming language	*Python*	*Python*	*Java/Python*
Neural network models	+	–	+
Vector of integrated words	+	–	–
Multilingual model	+	+	+
Tokenization	+	+	+
POS tagging	+	+	+
Segmentation	+	+	+
Parsing	+	–	+
Highlighting named objects	+	+	+
Communication between objects	–	–	–

Having studied the technical possibilities for the implementation of the semantic analysis algorithm and training the neural network, the authors will use the Spacy and StanfordNER libraries. The StanfordNER and Spacy libraries allow us to model our own model. It also allows you to make

the necessary configurations, depending on the specifics of the (Kazakh) language in question.

It is necessary to define the StanfordCoreNLP settings [15]: token – tokenize; ssplit – distribution of offers; pos – speech definition; lemma – find the original form of each word; ner – highlighting named objects; – regexner – work with regular expressions; parse – semantic analysis of each word; depparse – definition of syntax between words and sentences.

Further, Figure 1 shows the developed algorithm for the implementation of semantic analysis taking into account keywords and describes the work of the modules.

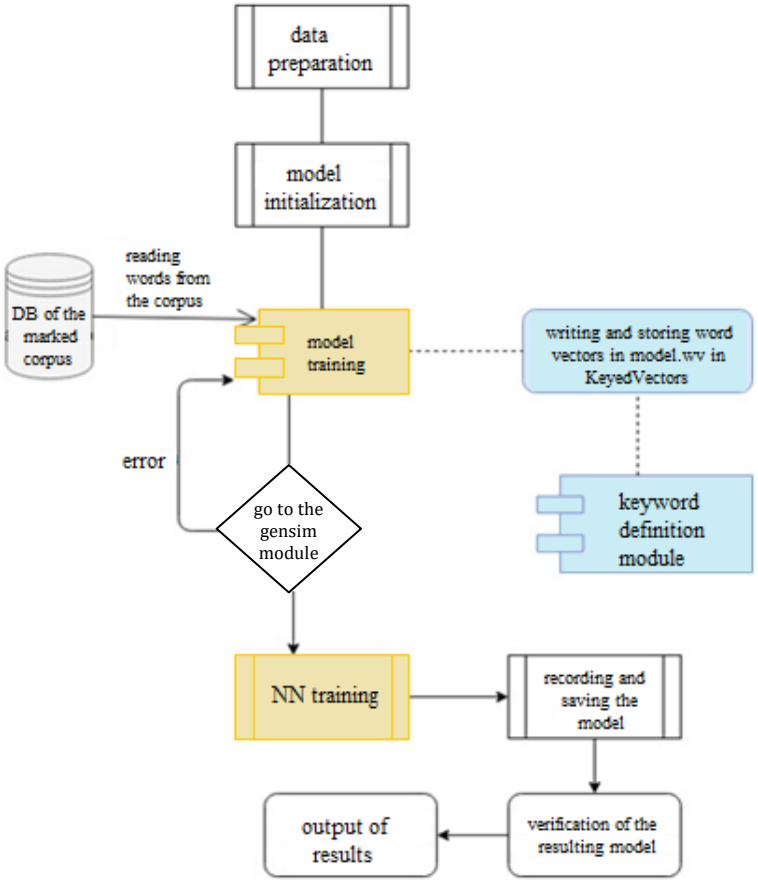


Fig. 1. Algorithm for the implementation of semantic analysis taking into account keywords

The input is text data. To train the model, set the following parameters: the dimension of the feature vectors is 100; the maximum distance between

the current and predicted word in a sentence is 5; the minimum education level is 1; the cutoff frequency is 4 words.

```
>>> model = WordVec(sentences, size=100, window=5, min_count=5, workers=4)
```

Record initialized model

```
>>> model.save(fname)
```

```
>>> model = WordVec.load(fname) #
```

Now you can train with the resulting model. For training the model, a monolingual Kazakh corpus was prepared, which is in the SQL database. When the text is processed by the model, vectors of words are identified, which are stored in the model.wv module in KeyedVectors. The resulting vectors of words are also compared with keywords (phrases) from the text corpus for the purpose of further use as possible values of semantic attributes of entities. Once the model finishes training, you can go to gensim.models.KeyedVectors in wv:

```
>>> word_vectors = model.wv
```

```
>>> del model
```

The gensim.models.phrases module automatically detects a long chain of words. This module allows us to define phrases through learning.

```
>>> bigram_transformer = gensim.models.Phrases(sentences)
```

```
>>> model = Word2Vec(bigram_transformer[sentences], size=100, ...)
class gensim.models.wordvec.Corpus(dirname)
class gensim.models.wordvec.LineSentence(source,max_sentence_length=10000, limit=None)
```

After completing the gensim module, you can then start training the neural network.

```
sentences = LineSentence('myfile.txt')
```

```
from gensim.models import Word2Vec # define training data
```

```
sentences = [['ұл (ul)', 'балалар (balalar)', 'қыздарға (qyzdarga)', 'қарағанда (qaraganda)', 'мықты (myqty)', 'болады (bolady)'],
['Ал (Al)', 'қыз (qyz)', 'балалар (balalar)', 'ұлдарға (uldarga)', 'қарағанда (qaraganda)','нәзік (nazik)'],
['Қыз (Qyz)', 'әлемнің (alemning)', 'көркі (korki)'],
['Гүл (Gul)', 'жердің (zherding)', 'көркі (korki)'],
['Қазақстан (Kazakhstan)', 'республикасы (respublikasy)', 'тәуелсіз (tauelsiz)', 'мемлекет (memleket)']]…
```

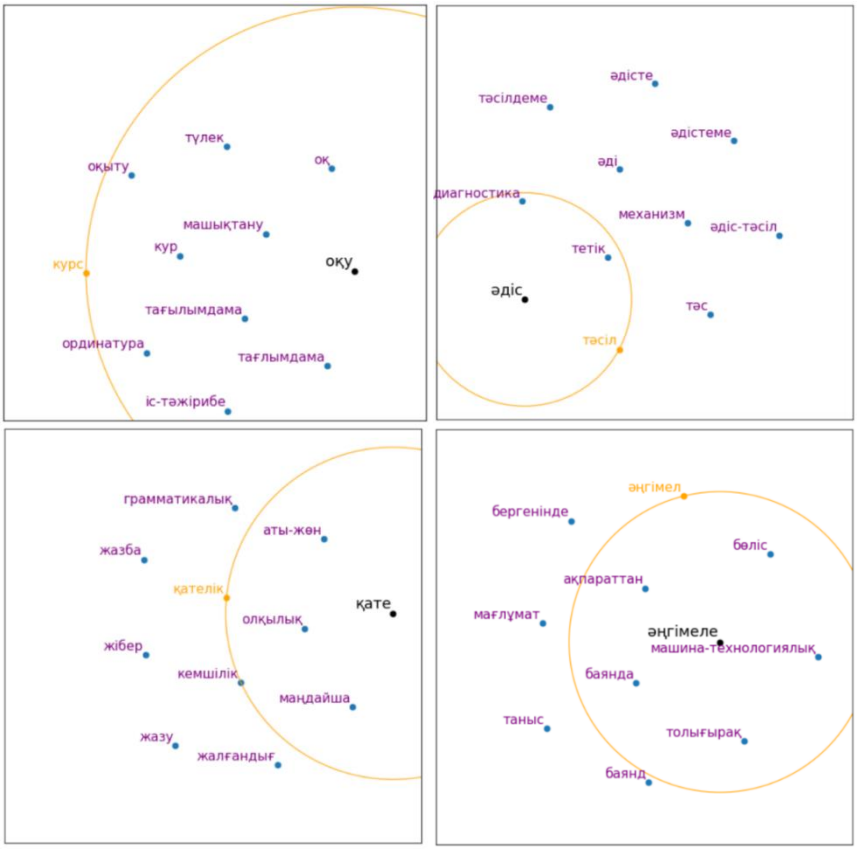
As a result of the obtained trained model, it is necessary to check the obtained data. You can also create a graphical interpretation of the results (Figure 2).

The NER Stanford software package was used to train the model. The following is a listing of working with the Stanford NER library and software implementation

```
>>> trainFile = train/dummy-kazakh-corpus.tsv serializeTo
= dummy-ner-kazakh-french.ser.gz map = word=0,answer=1

useClassFeature=true useWord=true useNGrams=true
noMidNGrams=true maxNGramLeng=6 usePrev=true useNext=true

useSequences=true usePrevSequences=true maxLeft=1
useTypeSeqs=true useTypeSeqs2=true useTypeySequences=true
wordShape=chris2useLC useDisjunctive= true
```



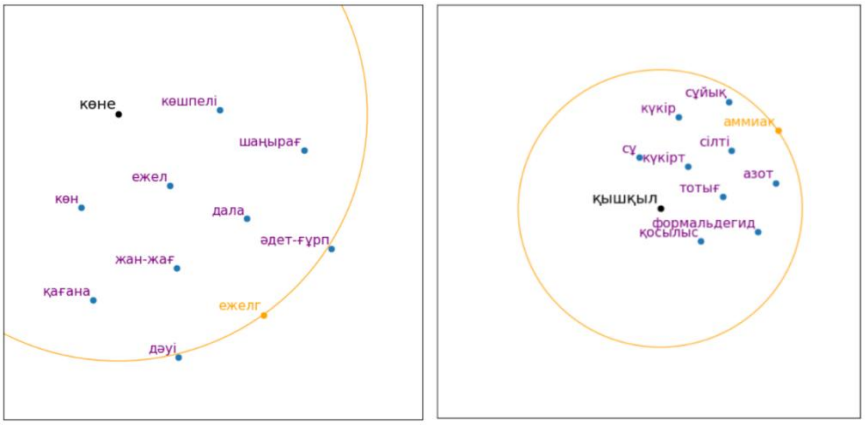


Fig. 2. Graphical representation of the vector space of practical results of the semantic analysis of the text in the Kazakh language

```
>>> cd stanford-ner-tagger/

java -cp "stanford-ner.jar:lib/*" -mx4g
edu.stanford.nlp.ie.crf.CRFClassifier -prop
train/prop.txt
```

```
# coding: utf-8

import nltk
from nltk.tag.stanford import StanfordNERTagger

# Optional
import os
java_path = "C:\Program Files (x86)\Java\jdk1.8.0_201"
os.environ['JAVA_HOME'] = java_path

sentence = u"Қазақстанда алма өседі. Алматы қаласында ҚазҰУ жоғары оқу орны орналасқан"

jar = './stanford-ner-tagger/stanford-ner.jar'
model = './stanford-ner-tagger/my-ner-model-french.ser.g'

ner_tagger = StanfordNERTagger(model, jar, encoding='utf8')

words = nltk.word_tokenize(sentence)
print(ner_tagger.tag(words))
```

Fig. 3. An example of input data of text for the program NER

The problem was successfully solved by using a morphological parser for marking up parts of speech in texts with the subsequent application of the machine learning method of semantically related keywords (phrases). A trained neural network with a hidden layer is applied to the set of these phrases in order

to assign a specific phrase to a specific attribute of the entity described in the text. Thus, based on a set of semantically related pairs of words, an ontology is built for a specific document, which is formed during the operation of a neural network.

5. Conclusions and future work

Methods and modern approaches to semantic analysis of texts have been investigated. An approach to the semantic analysis of texts in the Kazakh language has been developed; The developed approaches and algorithms were applied for processing texts in the Kazakh language; To solve the problem, semantic analysis in the Kazakh language is based on machine learning. The program is implemented in the python programming language, using the libraries gensim, matplotlib, sklearn, numpy, etc. A set of vectors of words in the Kazakh language was obtained, which was trained on the corpus, which is one million sentences. The corpus is fed to the program input in a normalized form. Further, to improve the result, the corpus will be supplemented with proposals on various topics. This developed approach will be applied in the development of the full post-editing module of the Kazakh language in the machine translation system and other applied tasks for natural language processing.

Acknowledgments

This research performed and financed by the grant Project IRN AP08052421 Ministry of Education and Science of the Republic of Kazakhstan, Project title: Research and development of the post-editing system of the Kazakh language in machine translation, by Research Institute of Mathematics and Mechanics, Al-Farabi Kazakh National University.

Bibliography

[1] Pospelov D.A., Ten hotspots in artificial intelligence research, Intelligent Systems (MSU), 1996, 1, 1–4: 47–56.
[2] Alypansky G.A., Braslavsky P.I., Titov P.V., Formation of information queries to Internet search engines based on the thesaurus: a semantically oriented approach, Proceedings of the VIII Intern. conf. on electronic publications EL-Pub2003, 2003, 269–270.
[3] Semantics, http://semantick.ru/. Last accessed: 14.07.2019.
[4] Tomita parser, http://api.yandex.ru/tomita/. Last accessed: 14.07.2019.
[5] In the foothills of semantics, http://dworq.com. Last accessed: 29.05.2019.
[6] AI Data Analysis Technologies for Business, https://www.summarizebot.com/summarization_business.html. Last accessed: 27.05.2019.
[7] TextAnalyst ver. 2.0 – Program for personal analysis of texts, http://offext.ru/library/data/datakeeping/51.aspx. Last accessed: 19.04.2019.

[8] Galaktika-Zoom – analytical system for respectable clients, https://www.itweek.ru/themes/detail.php?ID=52215. Last accessed: 16.06.2019.
[9] Automatic text analysis technologies, http://nlp.isa.ru/. Last accessed: 26.04.2019.
[10] GitHub natasha, https://github.com/natasha. Last accessed: 26.04.2019.
[11] Manning Ch.D., Raghavan P., Schütze H., Introduction to Information Retrieval, 2008.
[12] Efficient Estimation of Word Representations in Vector Space, https://arxiv.org/pdf/1301.3781.pdf. Last accessed: 10.07.2018.
[13] Word2vec Parameter Learning Explained, https://arxiv.org/pdf/1411.2738.pdf. Last accessed: 10.07.2018.
[14] Texts in, Meaning out: neural language Models in semantic similarity tasks for Russian, https://arxiv.org/ftp/arxiv/papers/1504/1504.08183.pdf: Last accessed: 20.04.2018.
[15] The Stanford Natural Language Processing Group, http://nlp.stanford.edu/software/CRF-NER.html. Last accessed: 20.04.2018.
[16] Erkan G., Radev D.R., Lexrank: Graph-based lexical centrality as salience in text summarization, Journal of Artificial Intelligence Research, 2004, 22: 457–479.
[17] Mihalcea R., Tarau P., Textrank: Bringing order into texts. Association for Computational Linguistics, 2004.
[18] Parveen D., Strube M., Integrating importance, non-redundancy and coherence in graph-based extractive summarization. In: proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015), 2015, 1298–1304.
[19] Khatri C., Singh, G., Parikh N., Abstractive and Extractive Text Summarization using Document Context Vector and Recurrent NeuralNetworks, 2018, http://arxiv.org/abs/1807.08000. Last accessed: 20.04.2018.
[20] Rakhimova D., Turganbayeva A., Auto-abstracting of texts in the Kazakh language, Proceedings of the 6th International Conference on Engineering & MIS, 2020, 1–5, https://doi.org/10.1145/3410352.3410832.
[21] Rakhimova D., Shormakova A., Problems of semantics of words of the Kazakh Language in the information retrieval, Proceedings of the 11-th International Conference 2019, Part II, 2019, 70–81.

Kunduz T. Sharsheeva[1], Gulnaz U. Tultemirova[2]

The study of methods for optimizing the route of movement in the city

Abstract: The problem of transport interchanges has been a big problem in many countries for many years. Every day the number of cars increases, and the roads were designed for a different number. Therefore, drivers spend several hours on the road every day. To optimize their route, drivers turn to the help of various applications, each of which has its own advantages and disadvantages. This article provides an overview and analysis of route optimization methods that can be used to find optimal routes around the city. Optimization methods for both hard (non-intelligent) methods and soft (intelligent) methods are considered, their comparative analysis and suitability for use, taking into account the features for a transport interchange. As algorithms for hard methods, the methods of hard computing and the method of compressed hierarchies are given, which use deterministic reasoning, clear classification, and binary logic to ensure accuracy, certainty, and rigor. These methods are widely described in the textbook Alekseev V. E., Talanova V. A. "Graphs. Models of computing. Data structures".

Fuzzy logic, genetic algorithms, expert systems, and neural networks are considered as algorithms of soft methods. Such methods differ from the analytical approach in that they use computational methods that are capable of representing uncertainty, vague concepts, and inaccuracies. The implementation of these methods in solving the problem of route optimization is called intelligent route optimization. Of course, in the modern world, intelligent methods that allow systems to learn, accumulate new knowledge, and adapt to new conditions are more applicable and developing. Therefore, special attention is paid to agent-oriented expert systems that allow agents to dynamically enter and exit the system, which are interconnected with each other to achieve individual or common goals.

Keywords: route, route optimization, hard computing methods, Dijkstra algorithm, fuzzy logic, expert systems, genetic algorithms, agent-based programming

1. Introduction

With the increased volume of vehicles, the problem of transport interchange. The situation is especially aggravated by the fact that many roads were built a long time ago, taking into account the load of the Soviet era. Now residents of the country have cars that create large congestion in Bishkek and impede the movement of drivers, who sometimes spend several hours moving from one point to another.

[1] Kunduz T. Sharsheeva, KSTU named after I. Razzakov, Ch. Aitmatov av. 66, Kyrgyzstan; e-mail: kunduz2000@mail.ru

[2] Gulnaz U. Tultemirova, KSTU named after I. Razzakov, Ch. Aitmatov av. 66, Kyrgyzstan; e-mail:gunya-t@mail.ru

Fig. 1. Traffic congestion in Bishkek

In front of people from one place, there is increased traffic on the route of movement, which as a result leads to congestion. This congestion affects the road traffic system and causes many negative impacts: increased transport costs, environmental pollution, etc. Therefore, nowadays users use effective planning in a dynamically changing environment.

Fig. 2. Air pollution in Bishkek

A range of navigation aids are available to assist drivers, but basic navigation systems consist mainly of analysis of static parameters (such as total route length or road type), and recent solutions rarely use vague data. In modern navigation

systems, the dynamic properties of traffic conditions are not yet sufficiently taken into account. For example, Google Maps only partially takes into account current traffic information. In congested cities, route planning is especially important in the day-to-day operations of public services such as police, emergency medical services, etc.

2. Overview of existing systems for building a passage

1. **Google Maps** is a set of applications built on the basis of a free mapping service provided by Google. The service is a map and satellite images of the planet Earth. Developed in 2005 by the developers Eilstrup Jens and Lars Rasmussen.

The functionality of the application:

- plotting a route from point A to point B;
- google Street View-view pre-captured photos in 360 degrees;
- directory of establishments;
- track your location.

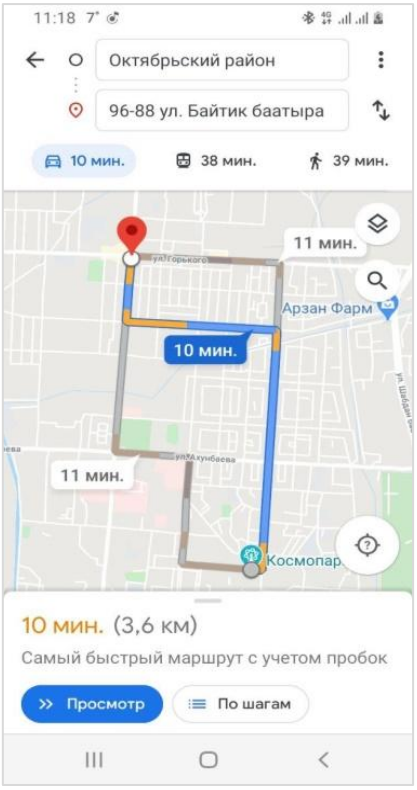


Fig. 3. Google Maps

2. **Yandex Maps** is a mapping service that has similar functionality to Google Maps, but with a number of functional differences. The map was developed in 2004 by Yandex. Unlike Google Maps, Yandex.Maps have a “Conversations” feature that allows you to leave comments on the map, essentially turning the app into a social network.

At the moment, the audience of the app has more than 50 million users worldwide, and this is only based on the mobile app.

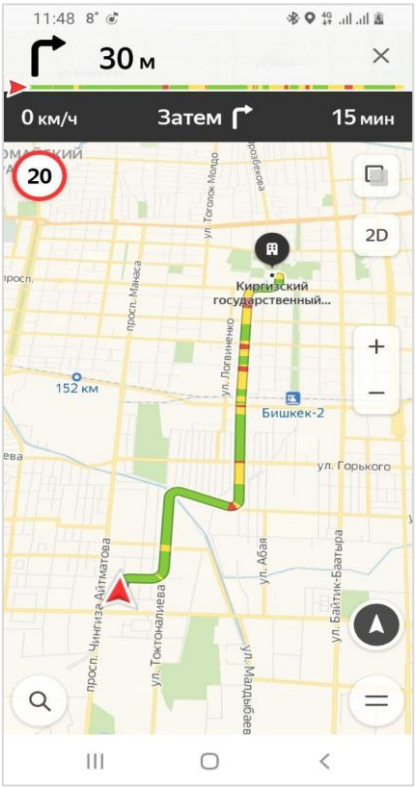


Fig. 4. Yandex Maps

3. **2GIS**-produces electronic reference books with city maps of the same name since 1999. Unlike its similar applications, 2GIS features more detailed rendering and detailed information about establishments, companies, etc. on the map. It has the functionality of a directory of establishments in a particular city. There is also a function of detailed route construction on public transport.

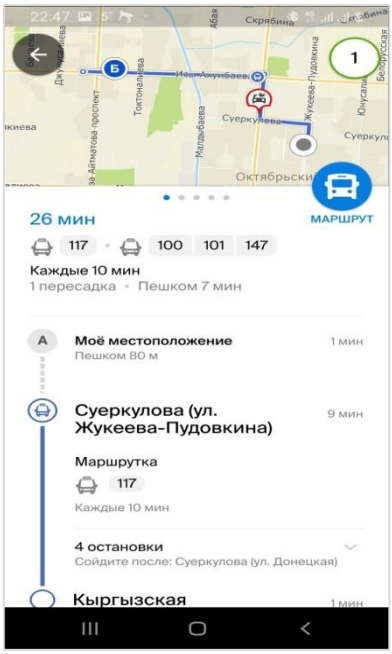


Fig. 5. 2GIS

Basically, the current systems for calculating traffic congestion use data on the number of vehicles per a certain section of road and the speed at which they move.

At the moment, the current systems of building a passage have shortcomings that lead to unrest and increased traffic jams on the roads.

For example, you can take a German artist who created virtual traffic jams by taking a cart with 99 mobile phones and walked around Berlin, creating a fake traffic jam. Based on this, we can conclude that the current system of building a passage is not ideal.

Subsequently, the car traffic fell on other streets, which in consequence created even more traffic jams and made it difficult to pass not only civilian vehicles, but also special services, which ultimately worsened the quality of public services.

3. Route optimization techniques

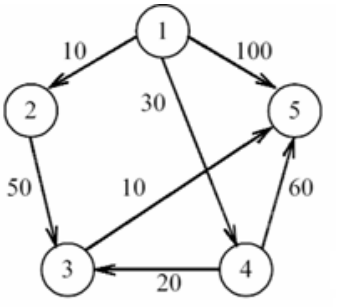
The road network is considered with reference to the theory of graphs as a graph of positive weights, the nodes of which correspond to the intersections of the road, and the edges of the graph represent the sections of the road (path) between the intersections. The length (distance) of the road section represents the weight of the edge. Several algorithms take advantage of these properties and

are therefore capable of computing the shortest path faster than using generic graphs. These include hard computing techniques, soft computing techniques and the agent-based computing paradigm.

3.1. Hard computing methods

Hard computing methods are analytical approaches that use deterministic reasoning, clear classification, and binary logic to ensure accuracy, certainty, and rigor.

Dijkstra's algorithm. For example, when the basis of the algorithm for finding the shortest path is Dijkstra's algorithm. Dijkstra calculates the shortest paths from a particular source node to all other available nodes in the graph, preserving the time distances for each node.



Iteration	***S***	***w***	**D[2]**	**D[3]**	**D[4]**	**D[5]**
Start	{1}	–	10	∞	30	100
1	{1, 2}	2	10	60	30	100
2	{1, 2, 3}	3	10	50	30	90
3	{1, 2, 3, 4}	4	10	50	30	60
4	{1, 2, 3, 4, 5}	5	10	50	30	60

Array P		1	4	1	3

shortest way from 1 to 5 {1, 4, 3, 5}

Fig. 6. Example of a graph used in Dijkstra's algorithm

Nodes are visited in the order of the algorithm following the shortest path from the source. The loop stops after visiting all the nodes of the target. Dijkstra's algorithm solves the problem on the shortest path from a single source and is not suitable for graphs with negative edge weights. The Dijkstra search algorithm is an algorithm that is usually useful when traversing a graph for traversable paths involving multiple nodes. The algorithm also uses heuristics to improve performance over time.

Compressed hierarchies are methods that introduce labels into the network, ensuring that nodes are ordered by importance. Then, a hierarchy is created by iteratively compress the least significant of the site. Shortening a node p means changing the shortest paths that pass through p by using shortcuts. Compression hierarchies intuitively assign different “significance levels” to each node. After that, the nodes are compressed in the significance hierarchy, removing them from the graph and replacing the labels to protect the shortest path distances connecting the more significant nodes. The various methods mentioned above can be used in combination, resulting in a more skillful algorithm than using a single method.

3.2. Soft computing methods

Soft computing methods differ from the analytical approach in that they use computational methods that are capable of representing uncertainty, uncertain concepts, and inaccuracies. The implementation of these methods in solving the problem of route optimization is called intelligent route optimization. These methods include: fuzzy logic, genetic algorithms, expert systems,and neural networks [3].

Fuzzy logic is an extension of logical logic that can handle the idea of partial truth, that is, truth values between “completely true” and “completely false”. The primary ways of thinking of fuzzy logic are evaluation, not accuracy. Fuzzy logic came about as a result of natural human thinking, which is related to approximations, which makes it very important.

An artificial neural network (ANN), also known as a neural network, is responsible for **processing** information and is stimulated by the way the biological nervous system (brain) processes information.

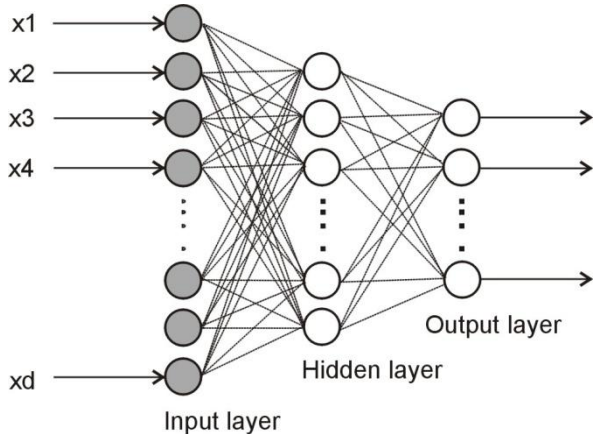


Fig. 7. Neural network

It consists of a large number of extremely unified processing elements (neurons) working in harmony to solve specific problems. Learning by example is one of the key features of ANNs, like people. In biological systems, learning involves fine-tuning the synaptic relationships that exist between neurons, this also applies to ANN. Nerve cells don't have to be the only system that can do neural computation, but an artificial system can also mimic the basic translation of a neural computing system. ANN is also known in various literature as parallel distributed processing, connection science, connectionism, and neural computing.

Genetic Algorithms (GA), which symbolizes a new programming paradigm that seeks to mimic the natural process of evolution when solving optimization and computational problems. In GA, sequences of bits, called computer **chromosomes**, are usually selected randomly from a set of computer chromosomes. This population is transformed into a new population by natural selection using operators stimulated by natural genetic operators. Inversion, crossover and mutation operators are operators identified by Holland that are used in the selection. The derivation of the fitness function is the basis for natural selection. Only a gene made with suitable chromosomes that survive can reproduce offspring. Between fit and less fit surviving chromosomes, more offspring are reproduced using suitable chromosomes than less suitable ones [4].

Natural selection operators function in this form:

1. The crossover operator. A trait (bit location) is selected from the parent descendants that crossover in a subsequence of a string before and after the selected location.

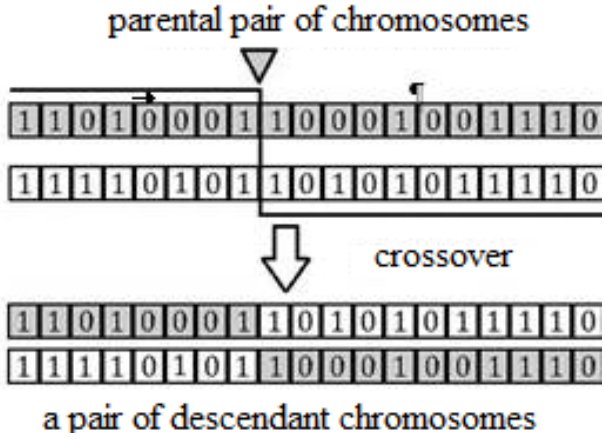


Fig. 8. Crossover operator

2. The mutation operator is the flipping of some bits in the chromosome.

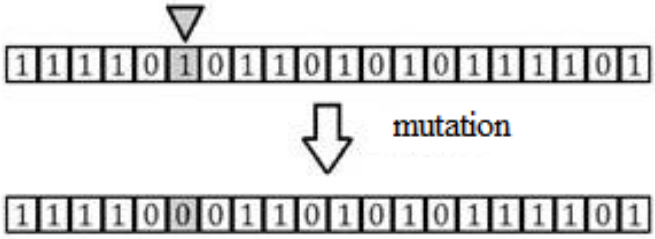


Fig. 9. Mutation operator

3. The inversion operator is to reverse the order of the subsequence of the chromosome.

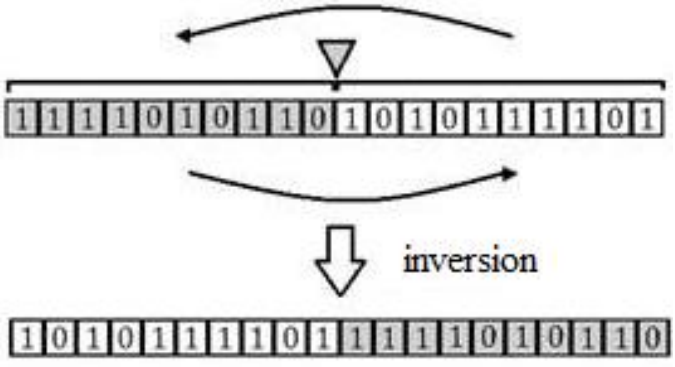


Fig. 10. Inversion operator

After the new generation of the population is complete, the evaluation criteria for stopping are met, and if it is, the algorithm stops. Otherwise, the fitness function is applied again to get the fitness rate of the new population. When solving a specific problem, the GA entry is an area of potential solutions, encoded by a metric fitness function that allows a quantitative assessment of each possible solution.

Expert systems (ES) are complex software systems that accumulate the knowledge of specialists in specific subject areas and replicate this empirical experience for the advice of less qualified users. Expert systems are a partial case of intelligent hypertext and natural language systems. Unlike conventional help systems, the user describes the problem, and the system concretizes it with

the help of an additional dialogue and itself carries out the flow of recommendations related to the situation.

Such systems belong to the class of knowledge dissemination systems. Control systems for mobile objects (in aviation, space technology, automobiles and other vehicles) are called semi-automatic control systems, when a person's ability to observe and assess situations arising during the movement of objects is used, and to form continuous control of them. One of the most developing areas of ES are agent-based systems [2].

The main method of forming judgments in the system is the construction of the reverse chain of inference. In this case, many meta – and object-level rules are used. As a result, it is very difficult to form a comprehensive and user-friendly explanation. The decisions that are made at the stage of programming the rules, in particular concerning the order and number of expressions in the antecedent part, can dramatically affect the path of searching in the space of solutions in the process of functioning of the system.

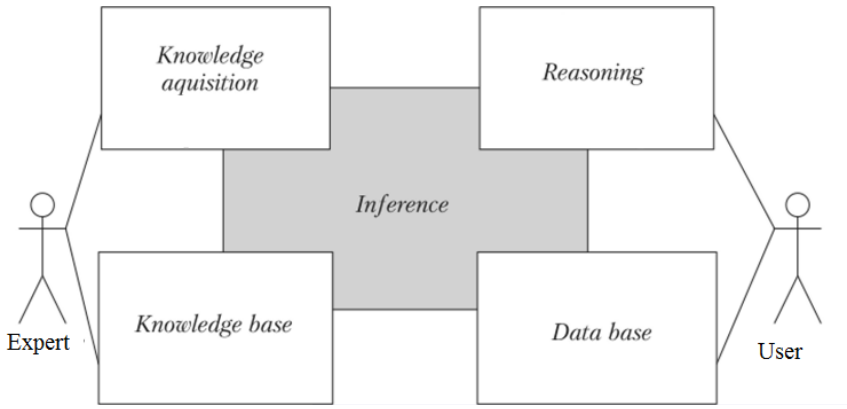


Fig. 11. Structure of expert systems

A significant part of the expertise of systems is the knowledge of typical cases, i.e., quite often found in the subject area. Experts easily recognize known model cases and are able to classify them in terms of ideal prototypes without much difficulty, even in the presence of certain "interference" or incomplete data. They intuitively distinguish between appropriate or unusual values of the source data and make an adequate decision about what to do in the future when solving the problem. Such knowledge is almost impossible to present in the expert system, if you use only the rules in the form of "condition-action". This will require a much more complex formalism, which will negate one of the main advantages of using generative rules as the main means of decision-making.

Agent-based programming (AOP) is a relatively new programming paradigm whose concepts have been transferred from theories of artificial intelligence to the traditional domain of distributed systems. AOP-based applications collect elements called agents, which are classified into categories: proactivity, autonomy, and communication. The autonomous nature means that separately they can carry out complex and often long-term tasks. Proactive means that proactive actions can be performed even without overt motivation from the user. Communicativeness means that interaction between other actors can take place to help achieve their own and other goals. The agent-centric architectural model of the application is basically peer-to-peer, where any agent is able to initiate communication between other agents. In agent computing, you can use more than one agent and agent role in the system that spawns a multi-agent system (MAC). MAS is a system of supportive or intelligent agents that interconnect with each other to achieve individual or common goals. As far as software development is concerned, one of the most important features of the MAC is that the last set of agents is usually not given at design time (only the first set is described). The implication of this is that traditionally, the architectures of multi-agent systems are open, which allows agents to dynamically enter and exit the system. The difference between an object-oriented approach and an agent in this sense is that objects can also log in and out of the system at runtime dynamically, but cannot do so autonomously as a result of proactive behavior.

Today, there are many options for solving the problem of optimizing traffic flow, both technical and organizational and managerial, such as: switching to small cars, creating new roads, increasing driving culture, using information technologies and mathematical methods for optimizing traffic, management models.

Optimizing road traffic with additional infrastructure will take years, a lot of money, and sometimes it's not even possible. For this reason, optimizing traffic signal timing is one of the fastest and most cost-effective ways to reduce traffic congestion at intersections and improve traffic flow in an urban network.

Researchers are working on using a variety of approaches to optimize traffic light timing. Use different solution methods. For example: traffic light tuning using Q-training and neural network approaches, optimization based on CI using a genetic algorithm or through artificial intelligence.

At present, the systems for optimal regulation of vehicle travel through intersections are insufficiently developed both in the Kyrgyz Republic and abroad. The principles of optimization when calculating the parameters of traffic light control are based on two basic approaches: computational (system-dynamic) and optimization (discrete-event). The calculation method is based on models of vehicle delay in front of traffic lights (models of M.J. Backmann, J.N. Darroch, W.R. McNeil, F.W. Webster, A.J. Miller, J.F. Newwell, etc.), according to which the control for the next cycle of traffic light operation is calculated based on the results of the passage of vehicles through the intersection

in the previous cycle. The optimization method uses a simulation model and an optimizer that take into account the current traffic flows, queues and crossing times in accordance with the optimality criterion. Both methods give significant errors [7].

The system-dynamic approach to transport modeling implies an increment of the model time with a constant step, regardless of the current network load. The calculation method is based on empirical formulas that make it possible to estimate the control parameters under which the best driving conditions will be provided. Such models are derived from the results of various studies and allow simulating the delay of vehicles at intersections under various operating conditions.

In the discrete-event approach, the movement of the model time occurs in jumps from one model event to another with the release of intermediate time intervals. At the same time, such intervals are omitted, during which no events occur that can lead to a change in the output data of the simulation. This approach consists in finding the best control parameters using optimization algorithms, and the parameters are estimated using a certain model that does not directly depend on the current time of traffic light control. The results obtained by the optimization method largely depend on the quality of the model used in the optimization loop.

Traditional mathematical models and analysis methods are not able to solve many real problems in transport technology, so modeling tools take a prominent place in the analysis of complex dynamic behavior of traffic flows.

Since the organization of traffic management is an area where conducting a full-scale experiment is difficult or impossible, simulation modeling in many cases becomes the only tool for effective decision-making in this area. One of the main advantages of this method is that, in contrast to the analytical one, the modeling of transport flows allows you to repeatedly reproduce the system under study and determine its optimal state. Thus, the creation of a simulation model of the city's transport network will allow you to clearly demonstrate the situation on the roads.

As a tool for simulation modeling, it is convenient to use the Anylogic environment – software for simulation modeling. Anylogic is the only simulation tool that supports all approaches to creating simulation models: process-oriented (discrete-event), system-dynamic, and agent-based, as well as any combination of them.

Another advantage of Anylogic is the ability to use machine learning. Reinforcement learning is a booming industry in machine learning. The construction of a fuzzy model is based on representing the characteristics of the system in terms of linguistic variables, which are considered as input and output variables of the control process. The fuzzy inference process is a procedure for obtaining fuzzy conclusions based on fuzzy prerequisites [5].

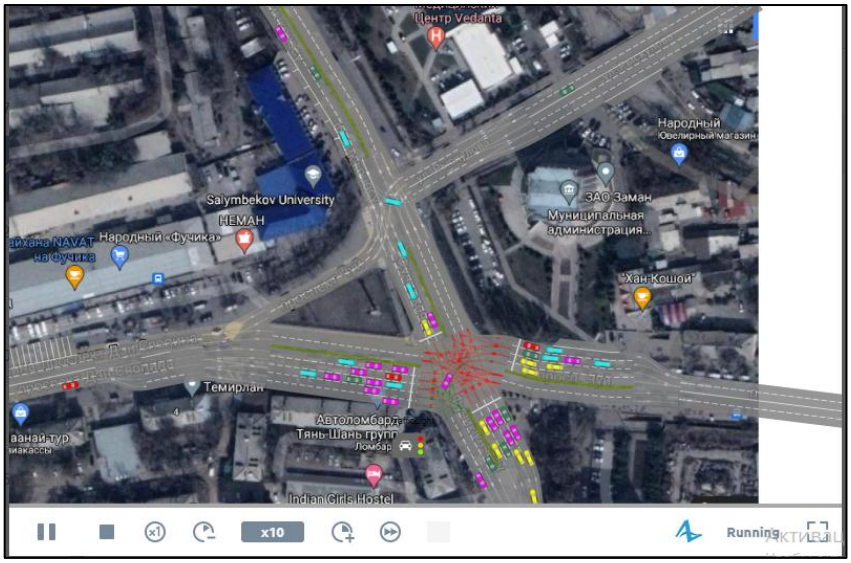


Fig. 12. An example of optimizing the operation of a traffic light at an intersection in Bishkek using a simulation model in Anylogic

The construction of an intelligent control system using this technology involves the following stages:

1. Determination of the inputs and outputs of the created control system.
2. Fuzzification (reduction to fuzziness) – assignment of membership functions for each of the input and output variables.
3. Formation of the base of rules of the fuzzy inference system.
4. Aggregation of subconditions in fuzzy production rules – determining the degree of truth of conditions for each of the rules of the fuzzy inference system.
5. Activation of sub-conclusions in fuzzy production rules – finding the degree of truth of each of the sub-conclusions of fuzzy production rules.
6. Accumulation of conclusions of fuzzy production rules – finding the membership function for each of the output linguistic variables.
7. De-diffusion of output variables – obtaining for each of the output linguistic variables of the usual (not fuzzy) quantitative value, suitable for use by external devices.
8. Checking the system operation.

As input parameters of the control system, various options for the characteristics of traffic flows at the approaches to the intersection can be considered: queue length, traffic delays, average spatial flow rate for a given period of time, etc., as well as their combinations, which is determined by the detectors used in a particular situation. transport. The output variables of the control process are the durations of the phases and the regulation cycle.

5. Conclusion

In the modern world, the most relevant solution to the problems considered is undoubtedly the development and use of intelligent systems of various levels. Such systems that use modern developments in the regulation of traffic flows (TP) and provide consumers with greater information content, due to the fact that they allow increasing the level of interaction of all traffic participants and regulate traffic optimization. Of course, in order to implement such optimal solutions for the design of the road network, it is necessary to take into account a wide range of different characteristics and the influence of various factors on the dynamic properties of these flows. But with the correct construction of the system, this is the cheapest option that road users can customize and use it.

Simulation modeling in many cases becomes the only tool for effective decision-making in this area. A tool such as modeling is very useful in traffic engineering to analyze the complex dynamic behavior of a traffic flow. Modeling can be defined as the imitation of real systems or processes to conveniently obtain information using similar traffic flow models. These models help describe the physical distribution of a traffic flow. The use of traffic simulation models is critical to comprehensively investigate the urban transport system in a safe and suitable environment.

Bibliography

[1] Ivanov V.M., I20 Intelligent systems: a tutorial, 2015: 92.

[2] Telkov A.Y., Expert systems unified collection of digital educational resources, 2017.

[3] Belikova T.P., Intellectual Systems, Handbook for laboratory work, 2011.

[4] Rana S., Examining the Role of Local Optima and Schema Processing in Genetic Search, 1999.

[5] Баязитов Г.А., Гибадуллин А.Р., Моделирование транспортных решений в среде anylogic, информационные технологии и системы, 2017.

[6] Фетисов В., Моделирование транспортных процессов, 2013, 164.

[7] Каталевский Д.Ю., Основы имитационного моделирования и системного анализа в управлении: учебное пособие, 2015.

[8] WebPromo, https://webpromoexperts.net/blog/vse-chto-vam-nuzhno-znat-o-reytinge-v-google-maps/. Last accessed: 24.07.2021.

[9] 2GIS, https://help.2gis.ru/. Last accessed: 24.07.2021.

[10] YMapsML, https://yandex.ru/dev/maps/ymapsml/doc/1.x/guide/concepts/usage.html. Last accessed: 22.08.2021.

Gulnaz Tultemirova [1], Kunduz Sharsheeva [2], Gulnara Oruzbaeva [3], Abdyldabek Akkozov[4]

Computer model of the holograms synthesis with a real phase

Abstract: Computer-synthesized holograms are widely used in areas such as optical information processing, image recognition, three-dimensional display of digital data, and modelling of holographic processes. It is difficult to overestimate the usefulness of the use of synthesized holograms for image reconstruction in acoustic and microwave holography. The use of synthesized holograms as elements of holographic storage devices is promising. Computer synthesis is often the only way to obtain holograms with desired properties. The main advantage of the synthesized hologram is that it is an effective means for converting digital information into optical. Due to this, it is possible to create hybrid-computing systems that are unique in performance, including digital electronic and optical processors and combining the flexibility and versatility of an electronic computer with the enormous performance inherent in an optical processor due to the parallelism of optical information processing. The use of digital holograms as elements in holographic storage devices is promising.

The authors of this work have developed an effective algorithm for phase shift refinement and a computer model of the processes of hologram synthesis with the real phase and optical restoration with the possibility of both qualitative and quantitative evaluation of the reconstructed image. Then, studies were carried out on this model, the results of which gave a positive answer that phase shift refinement using the correct phase method has a positive effect. The essence of this algorithm and the main points of research are briefly presented. Based on the modelling results, the dependences curves of the quality of the reconstructed image on the number of phase quantization levels were obtained for recording cases with the refinement of phase shifts by the method of real phases and without refinement.

Keywords: hologram synthesis, computer model, real phase, reconstructed image

1. Introduction

Digital holograms are used in various fields of science and technology, such as microscopy, interferometry, recognition of three-dimensional objects, in medical technology and are being introduced in the field of nanotechnology. It should be noted, also, an important feature of synthesized holograms, which

[1] Gulnaz Tultemirova, Kyrgyz State Technical University named after I. Razzakov, Ch. Aitmatov av. 66, Bishkek, Kyrgyz Republic; e-mail: tultemirova@gmail.com
[2] Kunduz Sharsheeva, Kyrgyz State Technical University named after I. Razzakov, Ch. Aitmatov av. 66, Bishkek, Kyrgyz Republic; e-mail: kunduz2000@mail.ru
[3] Gulnara Oruzbaeva, Kyrgyz State Technical University named after I. Razzakov, Ch. Aitmatov av. 66, Bishkek, Kyrgyz Republic ; e-mail: gul_talg@mail.ru
[4] Abdyldabek Akkozov, Kyrgyz State Technical University named after I. Razzakov, Ch. Aitmatov av. 66, Bishkek, Kyrgyz Republic; e-mail: abysh2012@mail.ru

consists in the possibility of obtaining optical wave fronts that do not physically exist, but are specified only by a mathematical description.

In the field of synthesis of digital holograms, the Lohmann method is known, which one of the first methods is. Quite a few works have been devoted to the study of this method. Despite this, there is an incompletely investigated version, called the “real phase” method [3], which, according to the authors, eliminates the phase error in image restoration.

The quality of the synthesized hologram depends on:

1. What coding method is used in its synthesis?
2. A successful choice of the values of the parameters of coding, sampling and quantization.
3. The perfection of technology for obtaining holograms: methods and devices for registering synthesized holograms, as well as materials and their photochemical processing (from the optimization of its modes, etc.).

Here, coding is understood as the representation (approximation) of the recorded holographic function, which is in the general case complex-valued, in the form of a non-negative function if it is required to synthesize a hologram on amplitude media, and in the form of a purely phase function, i.e., in the form of a complex-valued function, but with a constant module, if it is required to synthesize a hologram on the phase environments.

Of the above factors that determine the quality of the synthesized hologram, the coding method plays a primary role. This is explained by the fact that, no matter how perfect the devices used for hologram registration, technology and photochemical processing, as well as methods for optimizing the coding parameters, the quality of the obtained holograms will still be poor if the coding method used in the synthesis is not perfect. But, at the same time, it should be taken into account that for each specific coding method there is its own limit of the achievable quality.

In other words, no matter how perfect the other components that affect the quality of the holograms are, the quality of the obtained holograms cannot exceed this limit. Therefore, to further improve the quality of the synthesized hologram, a more perfect coding method is needed. It should also be noted that any coding method is always developed taking into account its feasibility at the existing level of achievements in the fields of science and technology used for the synthesis of holograms. That is, at each qualitatively new level of achievements in science and technology, new coding methods should be reviewed and developed and / or improved.

The discreteness and binary structure of the hologram are the reasons for the appearance of diffraction orders. Of interest is only that part of the reconstructed image that is in the area

$$|x| \le \frac{a_I}{2},\ |y| \le \frac{b_I}{2}.$$

It is described by the function

$$\tilde{h}(x,y) = rect(\Delta\xi x) rect(\Delta\eta y) h(x,y). \quad (1)$$

The corresponding spatially limited discrete ideal hologram in the above-specified region reconstructs the image, which will be written in the form

$$\tilde{u}(x,y) = \Delta\xi\Delta\eta \mathrm{rect}(\Delta\xi x) rect(\Delta\eta y) \sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} U(n\Delta\xi, m\Delta\eta) \exp\left[i2\pi(n\Delta\xi x + m\Delta\eta y)\right]. \quad (2)$$

The synthesized hologram will have the necessary properties if (1) and (2) coincide with accuracy up to a constant factor:

$$\mathrm{conct}_1 \tilde{u}(x,y) = \tilde{h}(x,y). \quad (3)$$

Obviously, in the general case, for an arbitrary $\tilde{u}(x,y)$, no choice of parameters P_{nm} can be used to achieve identical coincidence of images in the required area. Therefore, it is more reasonable to require the best approximation in the sense of some criterion for the proximity of the desired images.

Consider the following approximation

$$\exp\left(i2\pi\Delta\xi P_{nm} x\right) \approx 1.$$

Neglecting this factor will result in distortion equivalent to the distortion caused by the presence of phase error in the aperture. The phase error $\delta\varphi_{nm} 2\pi P_{nm}\Delta\xi x$ can have a stronger effect on the image quality. To establish the magnitude of the phase error, recall the range of variation of its components:

$$|P_{nm}| \le \frac{1}{2|Q|} \quad ; \quad |x| \le \frac{a_I}{2}.$$

Consequently,

$$|\delta\varphi_{nm}| \le \frac{\pi}{2},$$

which is equivalent to an optical path error ranging from 0 to $\frac{\lambda}{4}$. Hence it follows that the phase error lies within the Rayleigh criterion for wave aberrations. This means that the desired phase error does not cause a noticeable deterioration in image quality. However, as you know, the Rayleigh criterion is valid for wavefronts described by sufficiently smooth functions, therefore, there are cases when the phase error causes a serious deterioration in the image quality.

2. Real phase method coding

It follows from the above that it is necessary to take measures to reduce the phase error. The source of the phase error is the displacement (shift) of the aperture in the Fourier plane relative to the reference point, as a result of which the phase of the diffracted wave during image reconstruction is equal to the phase of the Fourier image at the reference point, and not the actual phase per aperture. That is, the light passing through the aperture located at the coordinate point $(n\Delta\xi + P_{nm}\Delta\xi, m\Delta\eta)$ carries the information contained in the coordinate point $(n\Delta\xi, m\Delta\eta)$.

Therefore, the authors of this work, from heuristic considerations, believe that the above-mentioned phase error can be eliminated if each aperture is positioned so that its position corresponds to the phase value of the Fourier transform at a given point, and not to the phase value at the reference point.

A graphic illustration of the essence of phase coding in the simple Loman method and the real phase method is shown in Figure 1. Here, to denote the dependence of the main value of the argument (reduced to an interval $[-\pi, \pi]$) of the Fourier transform $U(\xi, m\Delta\eta)$ on the relative (reduced) abscissa $\bar{\xi} = \xi / \Delta\xi$, the function (4) is introduced.

$$\Phi_U(\xi / \Delta\xi) \equiv \varphi(\xi, m\Delta\eta) = \arg U(\xi, m\Delta\eta) = \mathrm{mod}\left[2\pi, ArgU(\xi, m\Delta\eta)\right], \quad (4)$$

and to designate the dependence of the main value of the restoring wave argument $R(\xi, m\Delta\eta)$, the function (5) is introduced.

$$\Phi_R(\xi / \Delta\xi) \equiv \arg R(\xi, m\Delta\eta) = \mathrm{mod}\left[ArgR(\xi, m\Delta\eta), 2\pi\right] = \sum_{n=-N}^{N} \mathrm{rect}\left(\frac{\xi - n\Delta\xi}{\Delta\xi}\right)\frac{2\pi}{\Delta\xi}(\xi - n\Delta\xi), \quad (5)$$

where $\mathrm{mod}[\psi, 2\pi]$ means the value of the angle ψ modulo 2π. Consequently, they are discontinuous functions with discontinuity points of the first kind.

From Figure 1 it follows that the solution of equation (5) consists in finding the abscissas of the intersection points of the graphs of the functions $\Phi_U(\xi / \Delta\xi)$ and $\Phi_R(\xi / \Delta\xi)$. Brown and Lohmann in [3] used the cubic interpolation formula

of Newton constructed from four neighboring sampling points to find the desired position of the aperture. They claim that this gave a definite improvement in the reconstructed image. However, this method of determining the "correct" phase is not always acceptable, since any interpolation polynomial contains a remainder, which, in the case of a phase function in individual sections, may not be small at all. This is because functions $\Phi_U(\xi/\Delta\xi)$ and $\Phi_R(\xi/\Delta\xi)$ are discontinuous. They have break points of the first kind. Breakpoints of a function $\Phi_U(\xi/\Delta\xi)$ can be located at any point, and breakpoints of a function $\Phi_R(\xi/\Delta\xi)$ can be at points corresponding to the boundaries of neighboring cells. Figure 1a shows that the existence condition for the simple Loman's method is the intersection of a straight line, which is satisfied for any pairs and. And according to Figure 1b that the condition for the existence of a real phase for the method is the intersection of a straight line, which does not hold for any pairs and. Therefore, for some cells, the value may remain undefined (see cell area in Figure 1b).

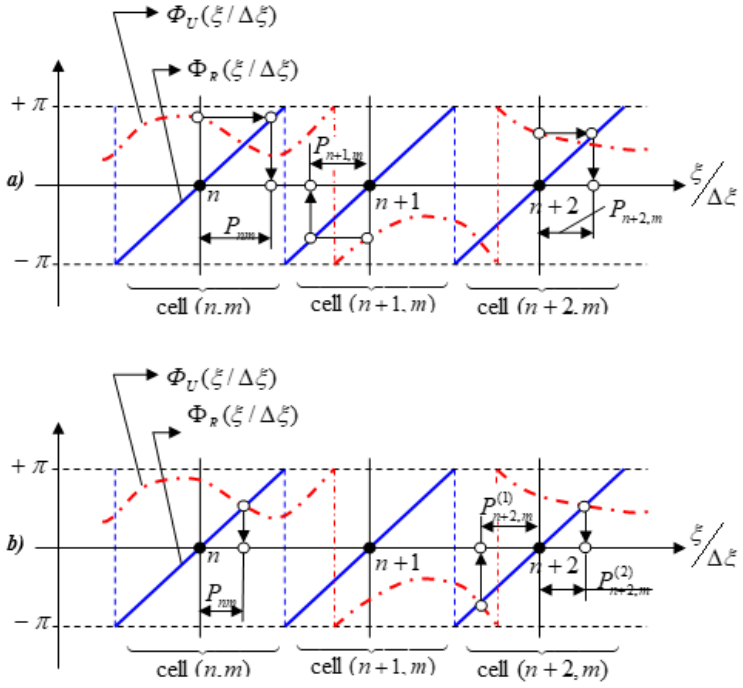
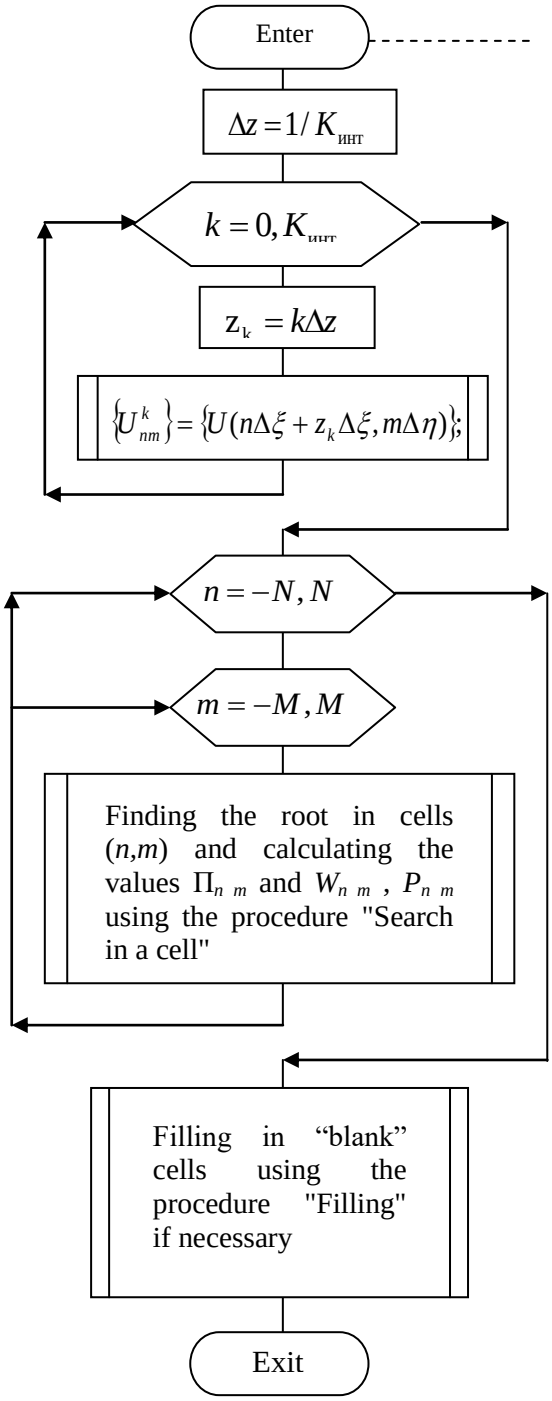


Fig. 1. Graphic illustration of the essence of phase coding by the Lohmann method (a) and the method of real phases (b)

Source: own elaboration



Input parameters:
$K_{ИНТ}$ – the number of subdivisions;
K_{pas} – gap accounting key $\psi(z)$;
$\Delta\xi$, $\Delta\eta$ – sampling steps;
N,M – numbers to determine the sample size $(2N+1)\times(2M+1)$.
Output parameters:
$\{W_{nm}\}$, $\{P_{nm}\}$ – an array of aperture sizes.

Fig. 2. Flowchart of the procedure “Method of the real phase 1”

A flowchart of the procedure “Method of real phases 1” is shown in Figure 2. The flowchart shows that after the completion of the calculations of the refined values P_{nm} and W_{nm}, as well as the values Π_{nm} for all cells, empty cells are filled, that is, cells with $\Pi_{nm} = 0$ if we want it. The point is that a set of empty cells, as will be shown later, can cause the appearance of bright light spots of small sizes in the centres of diffraction orders of the reconstructed image. If we want to eliminate these spots, the zero values P_{nm} and W_{nm} empty cells must be replaced with values calculated using simple coding ratios from the samples at the centres of the cells $U(n\Delta\xi, m\Delta\eta)$.

3. A study of the actual phase method by computer modelling

For a given holographic object, two holograms were synthesized: one of them is calculated with the phase shift refinement, and the other without refinement, i.e., using simple coding relations. In this case, the sampling and coding parameters are selected the same.

The quality of the reconstructed image is estimated by the root-mean-square-error (RMSE), i.e. the deviation of the reconstructed image from the original holographic object:

- by modulus of complex amplitudes;
- by complex amplitudes.

Figure 3 was chosen as the initial holographed object in computer modelling.

Fig. 3. Professor Abdyldabek Akkozov

Source: own elaboration

To clarify the dependence of the quality of the reconstructed image σ^2 on the degree of refinement K_{int}, the values of the root-mean-square error σ^2 were calculated for various values of K_{int}, from which the curves shown in Figure 4

were obtained. Curve 1 was obtained with a diffusely illuminated object. Curve 2 corresponds to the case of specular illumination of the object. It can be seen from this figure that it is sufficient to select the K_{int} value in the range from 4 to 8. An increase in K_{int} above practically does not lead to further improvement, but only leads to an increase in the calculation time. To exclude the influence of quantization of the encoded phases on the result, the synthesis of holograms was carried out without quantization (this is possible with computer simulation).

Based on the modeling results, the curves of the dependences of the quality of the reconstructed image on the number of phase quantization levels were obtained for recording cases with the refinement of phase shifts by the method of real phases and without refinement. Such curves were obtained both for a specular object (Figure 4) and for a diffuse one (Figure 4). In these figures, curve 1 corresponds to the case of synthesis without refinement, and curve 2 to the case of recording by the method of real phases. Comparing curves 1 and 2 it follows:

- phase refinement really gives a positive effect from the point of view of the minimum mean square error;
- the optimal value of the number of quantization levels is in the range from 16 to 48;
- the method of real phases is more effective for the case of the holograms synthesis of mirror objects than for the case of the holograms synthesis of diffuse objects.

For a qualitative (visual) assessment of the reconstructed image, pictures of the reconstructed image of a mirror object were obtained both for the case of synthesis by the method of real phases (Figure 6 (a)) and for the case of synthesis by the simple Loman's method (Figure 6 (b)). It follows from Figure 6 (a) that the method of real phases, although it gives a positive effect on the root mean square error, restores the image of the original object, on which unwanted bright light spots of small sizes are superimposed in the centres of useful diffraction orders. This phenomenon is a consequence of the fact that the cells with $\Pi_{nm} = 0$ remain "empty", that is, without apertures. In our example, with K_{int} = 16, the 128×128 of 1820 cells turned out to be empty. To eliminate these spots, the empty cells were "filled in", by replacing the zero values of the P_{nm} and W_{nm} of the empty cells with their values calculated using simple coding ratios based on the counts in the centers of the $U(n\Delta\xi, m\Delta\eta)$ cells, and also, the step of sampling $\Delta\xi$ is reduced by half. As a result, a reconstructed image is obtained, shown in Figure 6 (a).

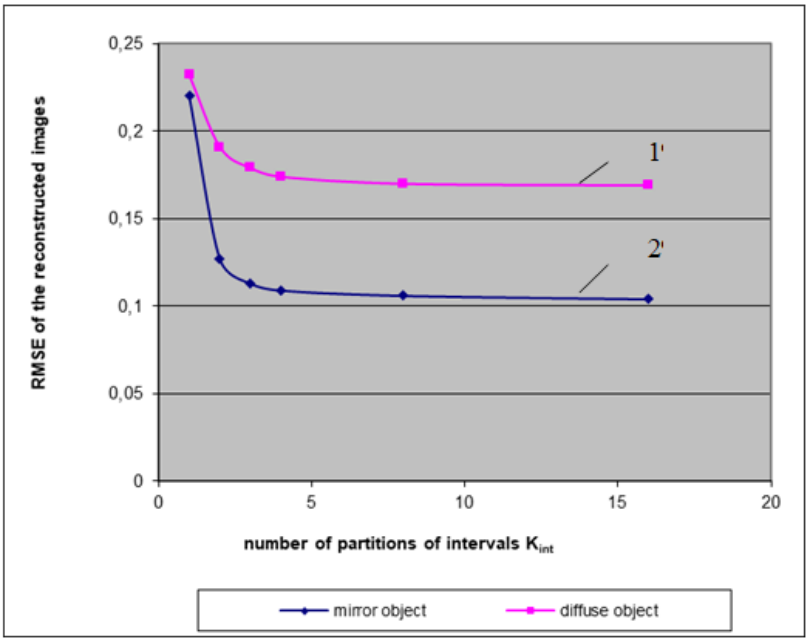


Fig. 4. Dependence of the root-mean-square error on the number of partitions Kint in the synthesis of a hologram by the real phases method

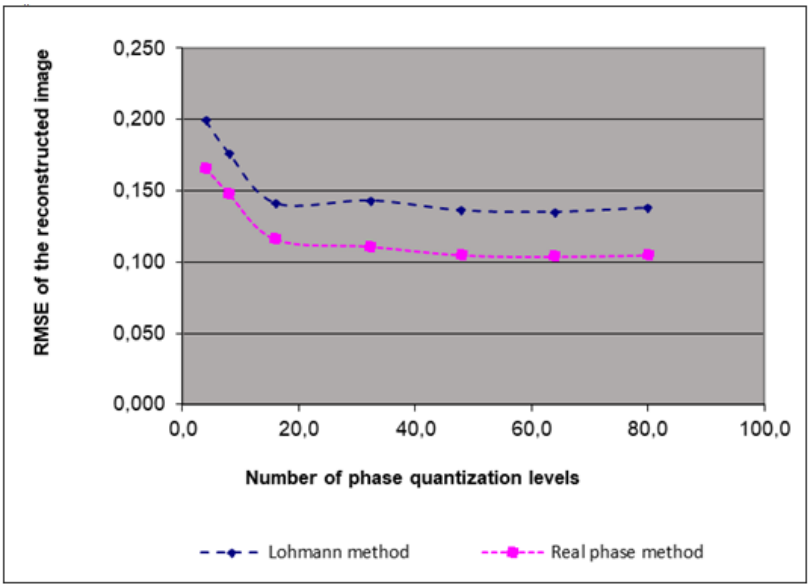


Fig. 5. Dependence of the root-mean-square error on the number of phase quantization levels in the synthesis of a hologram of a diffuse object

a)

b)

Fig. 6. The result of computer model of the holograms synthesis with a real phase:
a) Image reconstructed from a hologram synthesized by the real phase method.
b) Image reconstructed from a hologram synthesized by the simple Lohman method

Source: own elaboration

4. Conclusions

The process of image reconstruction was investigated with the help of cosmic computer modelling, considering the cases of matching and non-matching of the aperture shift with the real phase and comparing them in quality.

The effectiveness of this algorithm lies in the fact that:

1. The efficiency of the real phase method is retained for all values of the quantization levels numbers of amplitudes.
2. The problem is reduced to solving linear equations, the algorithm for solving which is "impossibly" simple.
3. The parallel solution of the equations for all cells reduces the number of necessary two-dimensional arrays for the simultaneous storage in the computer's RAM of the Fourier transform sample matrices at the nodes of the main and additional sampling grids to a minimum, namely, to two, regardless of the selected number of intermediate points.
4. Phase shift refinement by the real phase method has a positive effect.

The disadvantage of this algorithm is that:

1. At relatively low values, real phase method can create a memory problem.
2. The method of real phases although gives a positive effect on the root-mean-square error, restores the image of the original object, on which unwanted bright light spots of small sizes are superimposed in the centers of useful diffraction orders.

Bibliography

[1] Kostuk R.K., Holography: Principles and Applications, 2019, 133–151.

[2] Schnars U., Juptner W., Digital holography – digital hologram recording, numerical reconstruction, and related techniques, 2005.

[3] Bioucas-Dias J., Phase unwrapping via graph cuts, IEEE Trans. Image Proc., 2007, 16: 698–709.

[4] Brown B.R., Lohmann A.W., Computer-generated binary holograms. IBM Jorn. Res. Develop., 1969, 13(2): 160–168.

[5] Lohmann A.W., Paris D.P., Binazy Fraunlofer holograms generated by computer. Appl. Opt., 1967, 6(10): 1729–1748.

[6] Hugonin J.P., Chavel P.A., High quality computer holograms: the problem of phase representation, JOSA, 1976, 66(10): 986–996.

[7] Akkozov A.D., To the refinement of the phase shift in the synthesis of holograms by the Loman method, Izvestiya VUZ, Bishkek, 2008, 7–8: 190–196.

[8] Hennelly B., Sheridan J., Random phase and jigsaw encryption in the Fresnel domain, Optical Engineering, 2004.

[9] Nishchal N., Joseph J., Singh K., Fully phase encryption using fractional Fourier transform, Optical Engineering, 2003, 42: 1583–1588.

[10] Rivenson Y., A. Stern A., Javidi B., Improved depth resolution by single-exposure in-line compressive holography, Appl. Opt. 2013, 52: A223–A231.

Salamat Zhunusbayeva[1], Zhanna Alimzhanova[2]

Use of new technologies in cryptanalysis

Abstract: The problem of protecting information resources is currently becoming increasingly important. The report provides an overview of modern methods of cryptanalysis, namely the use of genetic algorithms as a tool for recognizing encryption algorithms. A number of quantum cryptanalysis algorithms are also analyzed. Their computational complexity is compared with the computational complexity of similar classical algorithms. It is concluded that when practical examples of a quantum computer appear, modern asymmetric encryption systems need to be improved and modified.

Keywords: neural networks, cryptanalysis, gene algorithm, ciphers, quantum computers, cryptography, cryptographic security

1. Introduction

This work is devoted to the most unusual and rather controversial approaches to the study of cryptographic systems. recently, there has been a sharp increase in the number of open papers on all issues of cryptology. Many cryptosystems, the stability of which did not cause much doubt, were successfully disclosed.

To date, many systems have been developed for which it has been proved that their stability is equivalent to the complexity of solving some important problems that are considered extremely difficult by almost everyone, such as the well-known problem of integer decomposition. Many of the disclosed cryptosystems were obtained as a result of weakening these supposedly persistent systems in order to achieve high performance. in addition, the results of extensive research conducted since the beginning of the XXI century, both in cryptography itself and in the general theory of computational complexity, allow the modern cryptanalyst to understand much more deeply what makes his systems unstable.

The emergence of new cryptographic algorithms leads to the development of ways to crack them. The result of each new method of cryptanalysis is a revision of the security assessments of ciphers, which in turn entails the need to create more reliable ciphers.

The history of cryptology – the science that combines cryptography and cryptanalysis-dates back many centuries, but this field of knowledge began to develop especially intensively with the advent of the computer era. In scientific cryptology, breaking a cipher does not necessarily mean discovering a method

[1] Salamat Zhunusbayeva, Al-Farabi Kazakh National University, Almaty, Kazakhstan; e-mail: ssdarhan@gmail.com

[2] Zhanna Alimzhanova, Al-Farabi Kazakh National University, Almaty, Kazakhstan; e-mail: zhannamen@mail.ru

that can be used in practice to recover plaintext from an intercepted encrypted message. Hacking is understood only as confirmation of the presence of a vulnerability of the cryptographic algorithm, indicating that the properties of the cipher do not correspond to the declared characteristics.

Conducting cryptanalysis for long-existing and recently appeared crypto algorithms is very important, since it is possible to say in time that this crypto algorithm is not stable, and to improve it or replace it with a new one.

In order to identify unstable crypto algorithms, it is necessary to constantly improve the already known methods of cryptanalysis and find new ones. there are quite a few publications devoted to neural networks as such, which consider their various types and architectures [1–4]. Purpose of this review study application of neural networks, genetic algorithms, quantum computers for cryptographic information security systems, as well as their application in cryptanalysis.

2. The need for cryptanalysis

The last decade has been characterized by a sharp increase in the number of open papers on all issues of cryptology, and cryptanalysis is becoming one of the most actively developing areas of research. Many cryptosystems, the stability of which did not cause much doubt, were successfully disclosed. At the same time, a large arsenal of mathematical methods of direct interest to the cryptanalyst was developed.

An ideal proof of the reliability of a certain public-key cryptosystem would be a proof that any algorithm for the disclosure of this system, which has a not negligible probability of its disclosure, is associated with an unacceptably large amount of calculations. Although no known public key system meets this criterion of strong durability, the situation should not be considered completely hopeless. Many systems have been developed for which it has been proved that their stability is equivalent to the complexity of solving some important problems that are almost universally considered extremely complex, such as the well-known integer decomposition problem. Many of the disclosed cryptosystems were obtained by weakening these supposedly resilient systems to achieve high performance. In addition, the results of extensive research conducted over the past ten years, both in cryptography itself and in the general theory of computational complexity, allow the modern cryptanalyst to understand much more deeply what makes his systems unstable.

Conducting cryptanalysis for long-existing and recently appeared crypto algorithms is very important, because over time, you can say that this crypto algorithm is unstable, and improve it or replace it with a new one. In order to identify unstable crypto algorithms, it is necessary to constantly improve the already known methods of cryptanalysis and find new ones.

3. The application of a neurocomputer network for cryptanalysis

In this method of cryptanalysis, input data is transmitted first:

- pain text;
- key;
- encryption algorithm;
- private text.

When the cryptanalysis process begins, only the closed text becomes known at first. Then, in the process of recognition, cryptotex can be represented as a single ensemble or as a single neuron. The response to the bite is represented by a response that the machine has learned from preliminary training.

In general, by this time, many articles have been published for the purpose of implementing a neural system in cryptanalysis, many of which have already been compiled into classical and current algorithms. The principle of their operation consisted of two goals. As for these goals, the first goal is to find a key based on existing Open Text and corresponding encrypted texts using pre-trained algorithm data, the second goal is to build a neural model for the studied cryptanalysis.

The basic principle of a neural system is to divide it into classes, i.e. divide it into classes. For its implementation, there are 3 mathematical models that:

1. Simple perseptron.
2. This is mainly intended for ideal linear tasks, which are basically represented as two different signals 0 and 1, now there are four possible points from these signals. And we can divide them into classes. These are (0,0), (1,1) – first class,(0,1), (1,0) – second class.
3. Multi-layer persepron.
4. These simple perseptrons create a giant multi-layer in large quantities, which is the alorithm that works in separation, finding an error in feedback.
5. Search for the largest number of matches, determining the weight of vectors in this algorithm.
6. To meet these mentioed requirements, there is a neuro-systematic model that has the opposite distribution. This model is the sum of two algorithms: the Kohonen layer and the Grossberg layer. The more complete an input vector is, the more complete it becomes. A schematic representation of the opposite distribution is shown in Figure 1.

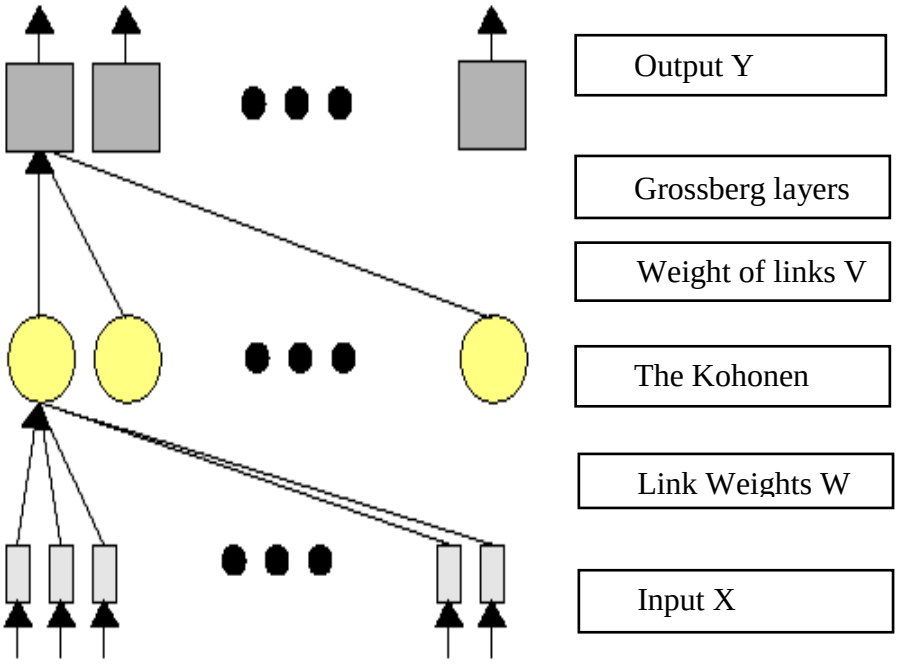


Fig. 1. Counter spread system

Now let's move on to this scheme: first the input vectors are denoted by 0 and 1, then their weight is determined, and it is denoted by W. Each neuron calculates its sum at the input in the Kohonen layer, and next to it, depending on the neurons in that layer, it determines whether it is active or not, that is, whether it is at level 0.1 or at level 0.0. According to this information, in the Kohonen layer, there is mutual competition between neurons, as a result of which an active neuron or a set of not very active neurons remain from the neurons, this layer can be considered as a filter. In this layer, a lot of calculations are performed, but thanks to this, active neurons can establish a connection between themselves and determine its weight, and, as noted, inactive neurons do not move to the next layer.

Next, the weight matrix of the neuron identified as active during training is determined, which is shown in the following Formula 1.

$$W_н = W_c + \alpha\,(X - W_c), \tag{1}$$

where $W_н$ – new value for neuron weight, W_c and this is the old essence, α the speed of learning, X – input vector length.

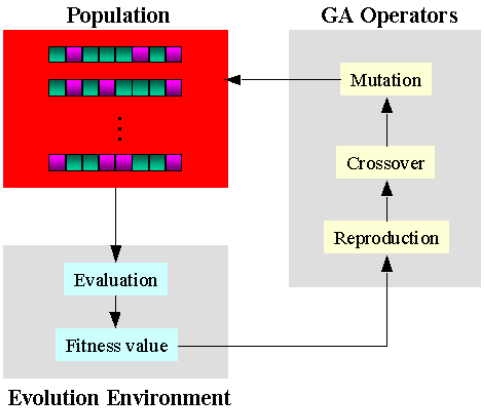
Learning algorithm: input: learning model (set of input vectors); output: adjusted relationships:

- specify the lines of the input vector;
- iterate until the constant state is set;
- adjust the connection according to the distribution of input data by hyperpersons for all nodes of the network or search for the largest match;
- repeat Steps 1–3 for each input vector.

4. Genetic algorithms

The genetic algorithm was first proposed by John Colmend. In this section of medicine, I present the information obtained as a result of research on this algorithm by John Colmend. In general, John Colmend's idea is to present a new documentary based on the input data, choosing independently. At the beginning of the work, a new population is created according to the algorithm, then a new population is created on the basis of a mutation, then a new population is formed on the basis of a mutation, and the original population turns into an old one.

In cryptanalysis, John Colland's algorithm is first presented as a binary string. Each binary is given as one chromosome. The very first step is to choose. As a result of selection, it is determined which new population will be the result of the selected chromosomes, and in this algorithm there is the concept of "parent", and it is chosen independently, but there is a greater chance of choosing the best. Then the algorithm will give you more chances that it will make the best choice [5]. Then their fusion begins, and the last stage is mutation. The specified process will continue to repeat until it leaves the loop. This process is shown in Figure 2 and Figure 3.



Genetic Algorithm Evolution Flow

Fig. 2. Evolution of populations

And in cryptanalysis, the genetic algorithm is developed into a “backpack” crypto system. That is, I considered it in this system. It is about the principle of the system, whether or not to introduce it into the mind [9]. That is, the question is which of the many things we put in the Rantz. For example, give us a set of values. $M_1, M_2, \ldots, M_n$ and their sum value S, then we need to find the value b_i.

$$S=b_1 M_1+ b_2 M_2+\ldots+ b_n M_n. \quad (2)$$

The value of b_I can only take two values, 0 or 1. If b_I = 1, then the *i*-th substance is placed on the rant, and if b=0, then the *i*-th substance is not placed on the rant. Then we can see the structure of the backpack in the form of chromosomes [7, 8].

The “best chromosome” selection function evaluates the proximity of the weight of a particular backpack to a given number. The function values are in the range [0, 1], where 1 means an exact match with the desired weight. If the weight of one backpack exceeds the target value *S* by some number *X*, and the weight of the other, on the contrary, is less than the required one by the same number *X*, then the last backpack is considered “best”. More formally, this function is described below:

1. Calculate the maximum discrepancy that can occur between an arbitrary chromosome and the target value of *S*:

$$\Delta_{max}= \max(S, \tilde{S} - S), \quad (3)$$

where $\tilde{S}$ is the sum of all the components that can be used when packing a backpack.

2. Calculate the weight of the backpack corresponding to the current chromosome, and denote S'

3. If $S' \leq S$, the”quality” of the chromosome is estimated by the value:

$$\alpha = 1 - \sqrt{\frac{|S'-S|}{S}}. \quad (4)$$

4. If $S' > S$, that $\alpha = 1 - \sqrt{\frac{|S'-S|}{\Delta_{max}}}$.

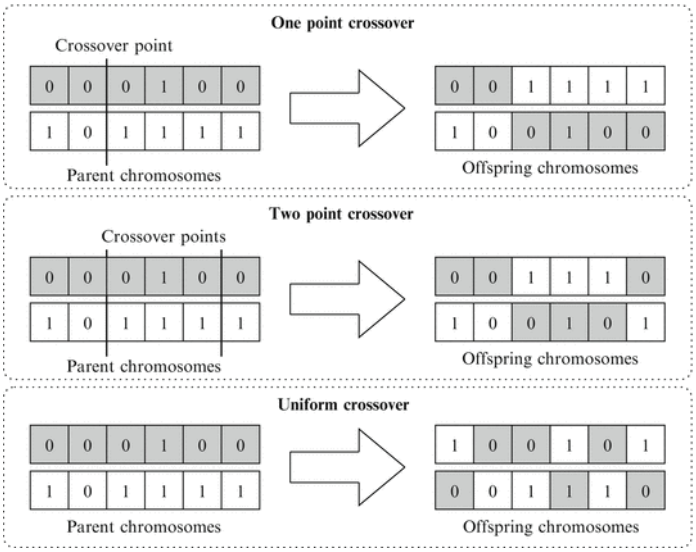


Fig. 3. Transmission of genetic information

After the mentioned process, the process of finding the best one within these chromosomes begins. The proximity weight is calculated to find the best chromosome. Let's show the algorithm as follows.

1. Binary chromosomes have a population and are performed randomly.
2. The evaluation function is considered for each chromosome.
3. According to the two steps above, natural selection is made in the form of the resulting coefficients.
4. In the above three steps, mergers are performed.
5. Off spring undergo mutations.
6. A new population is analyzed and the best chromosomes are selected.

This process is performed before the defined cycle, and the best chromosomes are used to decrypt [12]. As a result of the study, it became clear that the genetic algorithm is best used in cryptanalysis for methods of substitution and substitution. In the future, it is planned to use the genetic algorithm for symmetric digitization. In addition, it is planned to consider how the Fourier function can be used in this genetic algorithm, and analyze what image can be transmitted in its digitized value.

5. Quantum computers

In this review article, the focus was on new technologies, and among these technologies, the area that particularly attracts attention is that quantum computers calculate faster than conventional computers. Based on this basis, it is planned to consider cryptanalysis in quantum computers. Many algorithms have

been developed for quantum computers, one of which is the shore algorithm. Shore proposed his algorithm in 1994. Shore's algorithm was to perform certain calculations over a finite polynomial time [13,14]. The ability to create cryptotrading on the basis of quantum computers, which differs from other technologies in asymmetric algorithms. For example RSA, the El-Gamal algorithm is still weak in cryptotrading. The length of their keys in the past is very large. In addition to the shore algorithm, there is a Grover algorithm. It is this Grover algorithm that is best suited for cryptotrading. For example, if we make a complete count on a simple computer, we get the problem $O(2^m)$, where m is the length of the key. And in quantum computers, this difficulty would be many times less.

Grover considered the function $y = c(k, x)$ in his algorithm, where x is pain text, k is key, and y is private text. In addition, Grover considered two cases:

$$f(k) = \begin{cases} 1, & if\ c(k, x_1) = y_1 \\ 0, & if\ c(k, x_1) \neq y_1 \end{cases}. \tag{5}$$

You need to find the value of the argument at which this function is equal to 1.
1. Consider the following quantum algorithm.1. Bringing the quantum register to the state:

$$\frac{1}{\sqrt{2^m}} \sum_{t=0}^{2^m-1} |t\rangle . \tag{6}$$

2. Calculating the function f from this register:

$$\frac{1}{\sqrt{2^m}} \sum_{x=0}^{2^m-1} |t\rangle |f(t)\rangle . \tag{7}$$

3. Repeat $\frac{\pi}{4}\sqrt{2^m}$ times the procedure for increasing the amplitude of all t_i for which $f(t_i) = 1$. (This procedure is described below).
4. Measure the state of the register. The result will be equal to the desired key with a probability of about 2-n.
5. The checking of the result. If it is unreliable, the entire algorithm should be executed in one step.

The procedure for increasing the amplitude consists of two stages:
1. Change of amplitude from α_j *to* $-\alpha_j$ for all t_i such that $f(t_i) = 1$. This operation is a Z transformation over the last quantum bit of the register.
2. Inversion relative to the mean. This transformation can be written as follows:

$$\sum_i |t_i\rangle \ \rightarrow \ \sum_i (2\alpha_{\widetilde{n\delta}} - \alpha_i) |t_i\rangle , \tag{8}$$

where $\alpha_{n\eth}^{\sim}$ is the average amplitude.

The inversion with respect to the mean can be written as a matrix:

$$= \begin{pmatrix} \frac{2}{N}-1 & \frac{2}{N} & K & \frac{2}{N} \\ \frac{2}{N} & \frac{2}{N}-1 & L & \frac{2}{N} \\ L & L & L & L \\ \frac{2}{N} & \frac{2}{N} & L & \frac{2}{N}-1 \end{pmatrix}. \quad (9)$$

As Grover pointed out, if quantum computers were to emerge, the problem would reach $O(2^{n/2})$.

The algorithm mentioned above can be used to hack a hash function, but it should be noted that the bolk length should not exceed 256 bits [15].

As a result of research conducted for the purpose of review, we can call quantum computers the best technology. But given that quantum computers are not built in an iron state at the moment, it would be better to focus our attention on the genetic algorithm.

I would like to present interesting information to a quantum computer. World-famous physicists Jonathan Dowling and John Presnill on Twitter bet on the development of quantum computers in the next year. This dispute should end on March 1, 2030. If a precnim quantum computer is created, Dowling thinks that it is no.

Currently, a special team at Google is working on a quantum computer. When quantum computers are created, asymmetry algorithms will need to be improved.

6. Conclusions

The report provides an overview of modern methods of cryptanalysis, namely the use of genetic algorithms as a tool for recognizing encryption algorithms. The algorithm of quantum cryptanalysis was also analyzed. Their computational complexity is compared with the computational complexity of similar classical algorithms. it is concluded that when practical examples of a quantum computer appear, modern asymmetric encryption systems need to be improved and modified. After studying the latest software products, including the latest advances in machine learning, as well as the Shor algorithm and the genetic algorithm, we came to the conclusion that we will continue to research a genetic algorithm for cryptanalysis. In addition, a study of Grover's algorithm has been conducted, and work on this algorithm has been done in the past, because at the moment the quantum computer is not fully completed for cryptanalysis, so the review ended with the results of the study. The purpose of the review article was to identify the upcoming tasks, and in connection with the definition of these requirements, we came to the following conclusion, it is planned to

create a program based on a genetic algorithm that determines the behavior of the studied cryptographic ciphers.

References

[1] Gorbenko I.D., Cryptographic protection of information in information systems. Course of lectures, 2002.

[2] Schneier B., Applied cryptography, 2nd edition, http://nrjetix.com/r-and-d/lectures. Last accessed: 20.08.2021.

[3] Al-Ubaidy M.K.I., Black-box attack using neuro-identifier, Cryptologia, 2004.

[4] ANSI X3.92. American National Standard for Data Encryption Algorithm, 1981.

[5] Gladkov L.A., Kureychik V.V., Kureychik V.M., Genetic algorithms, Ed. Kureychik V.M., 2006.

[6] Avdoshin S.M., Savelyeva A.A., Cryptanalysis: the current state and prospects of development, 2012.

[7] Sharma L., Pathak B.K., Sharma N., Breaking of Simplified Data Encryption Standard Using Binary Particle Swarm Optimization, 2012.

[8] Sharma L., Pathak B.K., Sharma R., Breaking of Simplified Encryption Standard Using Genetic Algorithm, 2012.

[9] Chernyshev Y.O, Sergeev A.S., Dubrov E.O., Review of algorithms for Solving cryptanalysis problems based on bioinspired artificial intelligence technologies, 2005.

[10] Gorodilov A., Morozenko V., A genetic algorithm for determining the key length and decrypting a permutation cipher, 2008.

[11] Morozenko V.V., Eliseev G.O., Genetic algorithm for cryptanalysis of the Vizhiner cipher. Vestnik permskogo universiteta, Series: Mathematics, Mechanics, Computer Science, 2010, 1, 75–80.

[12] Sergeev A.S., On the possibility of using genetic search methods for implementing cryptanalysis of the asymmetric RSA data encryption algorithm, Izvestiya vuzov. Sev. Kavk. Region, Technical sciences, 2008, 3, 48–52.

[13] Kitaev A., Shen A., Vyaly M., Classical and quantum computing, 1999.

[14] Ozhigov Yu. I., Quantum computing, Educational and methodological manual, 2003.

[15] Saito A., Kioi K., Akagi Y., Hashizume N., Ohta K., Actual computational time-cost of the Quantum Fourier Transform in a quantum computer using nuclear spins, Quantum Physics, abstract quant-ph/0001113.

Part II

Data Analysis and Decision Making

Katarzyna Baran[1]

Thermal imaging of stress: a review

Abstract: In recent years, the subject of stress has become a key research area due to the increase in stress in people as a result of changing lifestyles, work pressure or the Covid-19 pandemic. Therefore, the introduction of preventive strategies for health protection, based on the latest technologies, can play a preventive and control role. The stress detection solutions used so far were often based on expensive devices for laboratory use, such as: ECG, EMG, GSR. A less expensive alternative is thermal imaging, which ensures non-invasive, comfortable examination, the possibility of using it in laboratory and home conditions, and which can use low-cost cameras while maintaining data detail. Combined with machine learning techniques, stress detection becomes more accurate and more reliable. Thanks to its features, thermography is more often used in interdisciplinary research. Therefore, this article focuses on stress thermal imaging, describing and comparing the latest research approaches. The most popular methods of stress detection, which are an alternative to thermography, are also presented. The most common ones are indicated: the analyzed physiological signals correlating with stress, the popular research devices used, questionnaires and classification algorithms. The limitations of thermographic stress detection are discussed, while at the same time pointing to selected research gaps in the field of stress recognition and detection.

Keywords: thermal imaging, detection stress, thermal camera

1. Introduction

The assessment of physical and mental stress has become the subject of research by many researchers, mainly due to the growing morbidity of the society, the occurrence of global and territorial pandemics and other factors causing the occurrence of increased stress. Long-term exposure to stress translates into damage to physical and mental health, causing, for example: depression, cardiovascular disease, heart disease [1 2, 3]. Monitoring and early detection of stress has a preventive and control power. Disease prevention is certainly less expensive than treatment afterwards. Especially now, when during the Covid-19 pandemic, a dramatic increase in the incidence of stress-related diseases is noticeable, which is associated with the need for diagnosis, treatment – and thus increases the need for medical assistance (e.g. psychologist, cardiologist) limited during the pandemic. All of this gives researchers an impetus to develop new techniques for recognizing and identifying stress.

So far, the most popular methods of detecting human stress have been the use of biological signals, especially: EMG, EEG, ECG, GSR. Despite the increasing efficiency of solutions and achieving high accuracy of stress recognition, the indicated methods often used expensive devices, not widely available, limiting

[1] Katarzyna Baran, Department of Computer Science, Lublin University of Technology, Nadbystrzycka 36B, 20-618 Lublin, Poland; e-mail: k.baran@pollub.pl

the comfort of the respondent (especially when it was necessary to put on/attach to the body), not ensuring complete non-invasiveness and causing additional stress. In addition, researchers [2, 3] noticed that among the commonly used methods, it is only possible to analyze a parameter or combinations of parameters, the values of which, unfortunately, do not provide absolute evidence of the accuracy of stress detection. With all this in mind, a search began for safer, non-invasive, portable and less costly solutions for stress testing. One of them is thermal imaging of stress, which is completely non-invasive and meets sanitary guidelines, especially tightened during a pandemic. Moreover, this method can take advantage of newer technological solutions and improvements in thermal imaging cameras.

Thermal imaging gained attention especially during the Covid-19 pandemic. A significant increase in the demand for thermal imaging cameras has been noticed. During the peak of the pandemic in Wuhan, drones equipped with thermal imaging cameras were used [4]. The American company FLIR has developed cameras that detect skin temperature up to 0.01°C [5]. However, it is worth bearing in mind that cameras can be deceived through, for example, additional emissivity sources. That is why companies are constantly developing automated and more precise systems. For example, the Polish company Scanway has introduced a camera to measure the temperature in the corners of the eyes and developed a system for remotely measuring the body temperature of an employee with the possibility of collecting data about their health [6]. Athena Security has launched a fever detection system that connects to a security camera system for real-time results [7]. Vodafone UK also introduced a camera equipped with IoT connectivity, with the ability to check the temperature of 100 people in a minute [8].

Thermography as an imaging process in the infrared range (wavelength from about 0.9 to 14μm), allows the registration of thermal radiation emitted by physical bodies in the temperature range encountered in everyday conditions, without the need to illuminate them with an external light source, and for accurate temperature measurement these objects. The temperature distribution on the surface of the human body depends on the temperature of the internal organs, the thermal conductivity of muscle and adipose tissue, and the emissivity of the skin. Any disturbances in the production of heat dissipation resulting from diseases of a given organ can be easily captured on a thermovision image of a human being. The thermal image of the diseased part or the whole organ will be significantly different from health imaging. Thermographic examinations are completely safe, non-invasive, and therefore are widely used mainly in medical fields and interdisciplinary work.

This article presents research approaches involving the use of thermal imaging for the recognition and classification of stress, as well as alternative stress detection solutions (the most popular), the most developed and refined by

researchers so far. Finally, the limitations of thermal research and research gaps in the field of stress detection are discussed.

2. Stress registration and human physiology

Stress, as the body's reaction to a threat, may manifest itself through changes in various human physiological signals. The most common in the literature include: temperature, blood volume pressure, galvanic skin reaction GSR, tidal volume, ECG, EMG, breathing patterns [3, 9, 10, 11, 12]. Changes in these signals can be caused by external factors (e.g. fear, emergency) or internal factors (e.g. pregnancy, a hot bath). An increase in any physiological parameter may also indicate stress. Skin temperature rises as stress levels decrease, and vice versa. Decreased BVP (blood volume pulse) values indicate increased stress and vice versa. During research, stress is most often caused by affective images, videos, texts, facial expressions, music or tasks [9]. Measurement devices for the indicated signals may require more careful processing to accurately identify the values. The following devices are used quite often [9]: *Emotiv EEG, temperature sensors, pulse converters, Simmer EMG, set of Biohames, Empatica E4, g.GSRsensor2, Tobii T60 Tracker*. Stress measurement is often supported by questionnaires [9]: *Standard Stress Scale*, *Ardell Wellness Stress Test*, *Stress Coping Resources Inventory*, *Perceived Stress Scale* and others. There are many problematic aspects in detecting stress, e.g. physiological variability, using developed data sets (e.g. DEAP, IAPS) instead of creating own, high costs of measuring devices. Looking at research approaches, the use of low-cost sensors and machine learning is increasing [3, 9]. Various classification algorithms are used in stress detection, with the most common: SVM [13, 14, 15], kNN [15, 16], random forest [14, 16], fuzzy logic [17].

During the detection of stress, physiological signals are characterized by variability of values correlating with the level of stress. When an EEG is used, a non-invasive bioelectric measurement of brain activity is performed using electrodes placed on the patient's head. The course of the recorded signals depends on the psychophysical state of the patient. The dominant rhythms in the recording of the EEG signal are called waves and are marked with the letters of the Greek alphabet (α, β, θ, δ, γ). The frequency of the signals varies in the range of 1–100Hz, and their amplitudes from several to several dozen μm. Too small share of alpha components is associated with hyperactivity and stress, and excessive – with concentration disorders. The purpose of treating mental disorders, addictions or stress therapy is neurofeedback – alpha training (8–13 Hz waves) allowing to achieve a state of complete relaxation.

When measuring body temperature under stress, the body "conserves resources" by inhibiting digestion. The circulatory system directs blood only to selected parts of the body. Stress hormones appear in the body: adrenaline, cortisol, and thyroxine. In the case of chronic stress, the body acts as it does in

the event of danger (despite its absence). In addition, psychogenic fever may occur with severe stress. Temperature control, relaxation, and the use of temperature biofeedback can restore normal conditions and eliminate the effects of stress.

The GSR (Galvanic Skin Response) technique is a method of measuring the electrodermal activity of the skin. In psychophysiology, it is one of the best-known measurements of the reaction to stress and experienced emotions. GSR sensors measure: skin conductivity, skin conductivity response and changes in sweating caused by increased or decreased levels of stress (sympathetic control). Stress is believed to stimulate the sweat gland, making the hands sticky and cold. GSR reflects the level of arousal caused by specific stimuli or messages. Larger values in the GSR notation mean greater amperage, i.e. less electrical resistance. The level of arousal of the cerebral cortex and skin conductivity are interrelated. The stimulation of the cerebral cortex is reflected in an increase in skin conductance, and a decrease in the level of arousal of the cerebral cortex is associated with a decrease in skin conductivity.

Measuring breath under stress involves measuring: CO_2 levels, chest expansion, and air volume. Human can control own breathing better than other organisms, therefore breath can be the basis for healing purposes. Slow lung function is able to bring into a state of relaxation, and stress is often associated with accelerated breathing to ensure a minimum of ventilation with decreasing respiratory capacity (stress affects the nervous system, mainly the hypothalamus and pituitary gland, mobilizing the body to react).

ECG, or electrocardiography, is a method of measuring the electrical activity of the heart muscle. The ECG waveform repeats three parts that correspond to the complete cycle of the heart: rushing of blood into the atria, contraction of the atria and ventricles, and pumping of blood out of the heart. The result of excessive stress can be a disturbance of the heart rhythm, called arrhythmia, and exceeding 60–100 beats per minute. If this condition persists for a long time, it can lead to serious changes in the body and even death. The analysis of HRV – heart rate variability is also important. Disease, aging or stress reduce this variability (greater involvement of the sympathetic nervous system), thereby increasing the risk of cardiovascular disease. Under stress, the heart rate signal is very uneven, there is low variability, only after focusing on the breath and equalizing it, the heart rhythm also becomes even. This is due to the fact that the rate of breathing directly affects the heart's rhythm. With HRV, stress levels can be found by analyzing the time intervals between heartbeats. Time distances are related to how stressed the patient is. The greater the variation in the time between strokes, the healthier the patient (the more involved the parasympathetic nervous system).

EMG, or electromyography, is based on the electrical activity of muscles. It allows to assess the functions of the muscular system and the peripheral nervous system. The EMG test can diagnose neuromuscular hyperactivity as well as diseases such as tetany. Due to the specificity of EMG, the test should not be

performed among people with implanted pacemakers, artificial valves or metal elements in or near the heart. The more stressed the patient is, the greater the electrical voltage in the muscle tissue. By giving the patient information about electrical activity from his muscles, the patient can learn to lower it (EMG biofeedback), which is often used in neuromuscular rehabilitation.

Figure 1 shows selected methods of stress detection, most commonly found in the literature, sample research tools and sample graphic interpretations of the recorded signals.

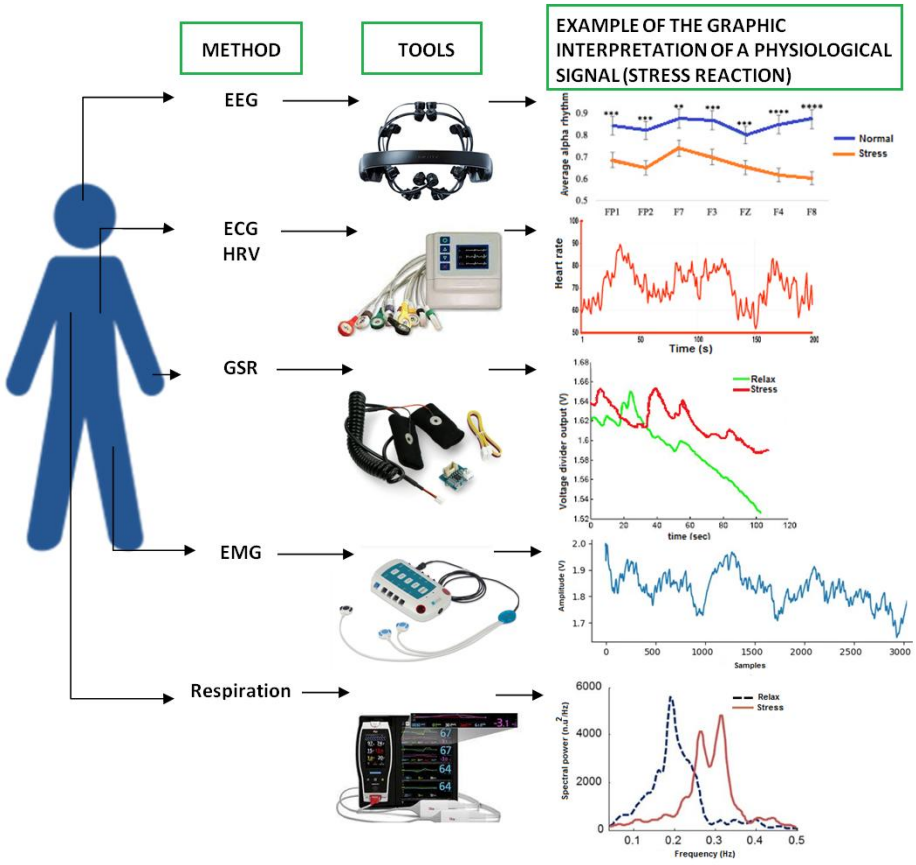


Fig. 1. Graphical presentation of selected methods of stress detection

Source: own elaboration

3. Related work – thermal imaging

Medicine and psychology often use advanced technologies that not only support diagnostics but also treatment. It is important that the technology is non-invasive and provides valuable, detailed information about the parameters of the subject correlating with the health. These criteria are met by thermal imaging, which is gaining attention especially in the Covid-19 pandemic. Selected recent research approaches using thermography in the field of stress detection are discussed below. While Table 1 presents a comparison of selected research approaches in the field of thermographic stress detection, taking into account also other research works. Figure 2 shows an example of thermal imaging of a person under light stress and a view of the FLIR thermal imaging camera.

Fig. 2. Thermal imaging: A) view of the subject's face under stress B) FLIR thermal imaging camera

Source: own elaboration

The stress load determined on the basis of simultaneous measurements of biosignals by means of an electrocardiogram (ECG), near infrared spectroscopy (NIRS) and nasal skin temperature (NST) in 10 subjects is presented in [14]. The study used the arithmetic mental task as a stress load. The ECG work monitor (CROSWELL Co., Ltd.) was used for the ECG measurement, the NIRS – OEG-SpO2 (Spectratech, Inc.), and the FLIR thermal imaging camera was used for the NST measurement. Feature extraction and biosignal selection were performed using the random forest method and the stepwise method (STEP). The SVM was used to estimate the stress. As a result of learning the model using the stepwise method, the F1 result was 0.765 (the highest result of all models). By analyzing

biosignal parameters, researchers found that the measurement of NST is more useful than ECG and NIRS in detecting stress during arithmetic tasks. NST decreases during task performance due to stress load, which was also confirmed by researchers in previous studies [10].

An interesting study that takes into account the correlation of stress with thermal changes in the face is presented in [18]. Thermal facial behavior and differences in vascular and inflammatory responses caused by acute social stress were analyzed in a study of 30 subjects. After each test according to the TSST protocol, interleukin-6 and mean blood pressure were measured and thermal images were captured using a FLIR A310 camera. Six regions of interest (biothermal biomarkers) were analyzed for thermal images: forehead, left and right cheek, chin, nose and corrugator muscle. Researchers found that in stressful situations, inflammatory activity and mean blood pressure increase – at different levels in men and women. Under stress, the direction and magnitude of thermal changes in the face depends on gender. The researchers also pointed out that there was no correlation between the inflammatory activity of the body and the peripheral temperature of the face.

Temperature patterns from the forehead and nose were used by researchers [19] to detect stress non-invasively. 9 subjects were filmed while performing TSST arithmetic tasks (Trier Social Stress Test). In addition to the ROI denomination, hierarchical GHSOM self-organizing maps, HGRBAC segmentation and Viola-Jones classification were used. Temperature in ROI areas was extracted as a function of time in order to identify patterns of stress build-up. The performance of each subject in TSST tasks was measured using the efficiency coefficient correlating with the level of concentration of the subject during the performance of tasks.

Automatic thermal stress detection in 45 people is also presented in [20]. Thermal (TS) and visible (VS) spectra were used, as well as the ANU StressDB database. TS films were recorded with a FLIR camera, and VS – with a Microsoft web camera. The face detection method based on the eye coordinates and the pattern matching algorithm (using MATLABs Template Matcher) was used in extracting facial areas. The Viola-Jones method was tested, but was inadequate due to the different nature of TS and VS films. Proposed a method for capturing dynamic thermal patterns in histograms (HDTP) to use thermal and space-time characteristics from videos. Patterns were compared with TS and VS video, and SVM was used for stress detection. The combination of HDTP patterns and functions with LBP-TOP allowed to achieve the highest stress recognition rate at the level of 72%.

In [15] the focus was on the detection of physical stress through hyperspectral imaging (HSI) of the saturation of facial tissues with oxygen (StO2). Twenty subjects analyzed the level of StO2, simultaneously determining five regions of interest (ROI): forehead, left and right cheek, nose, chin. SVM and 5-fold cross-validation were used for the classification. The combination of

selected ROIs allowed achieving the highest stress recognition rate at 82.11% (after removing the chin ROI). Five other algorithms were also tested, where the highest accuracy for the designated ROI was demonstrated by LD (Linear discriminant) – 80%. These researchers presented a similar study also in [13] examining the StO2 data in 42 subjects subjected to various types of stress (physical or mental). The PixelFly camera (PCO, Kelheim, Germany) was used together with the Specim VNIR spectrogram (SPECIM, SPECTRAL IMAGING LTD., Oulu, Finland). They indicated that tissue oxygen saturation values from the left cheek, chin and center of the eyebrow could provide the highest level of classification – 95.56% (SVM).

The two-level TSDNet stress detection network based on facial expressions and movement of a person was the idea of the researchers in [21]. The 2092 video dataset was the basis of TSDNet's performance testing. A detection accuracy of 85.42% was achieved with an F1-Score of 85.28%, and it was found that taking into account both facial expressions and human movements can improve stress detection by more than 7%.

Stress detection system (DeepBreath) using inexpensive infrared camera is proposed in [22]. The focus was on detecting respiratory signals, and their division was made by respiratory variability spectrograms (breathing without stress, with low stress, with high stress). The SCWT and Mental Computation tests were used to induce stress. In order to normalize the subjective ratings, the features were scaled based on the maximum and minimum scores. Using the CNN classifier, 84.59% classification accuracy was achieved for the binary classification and 56.52% for the 3-class classification. These researchers, a year later, presented a new *Thermsense* research system that uses the low-cost FLIR ONE camera to observe breathing under stress [23]. Also in [24], these researchers monitored stress by examining thermal variability of the nose through mobile thermal imaging outside of controlled laboratory conditions. And in [25], this team detected mental stress by combining PPG data from smartphone and thermal imaging, achieving the highest accuracy of 78.33% (PPG with thermal data).

The team [26] measured stress during surgical training in experienced and novices. The focus was on measuring responses to sweating in the periapical region. The discrete methodology of the researchers made it possible to compare the speed and dexterity of performing the task in the subjects and to confirm that the quick action of experienced surgeons results from skill, and in novices – from a high level of stress, which translates into accuracy.

Table 1. Comparison of selected research approaches

Study	**Tools**	**Stressors**	**Subject**	**Classifiers/ Algorithms**	**Best Accuracy Achieved**
[27] (2020)	MuSE dataset FLIR A40	QA – Question Answering, provocative video	28	RNN, CAs (Convolutional-Autoencoders)	59.9% (thermal modality)
[28] (2019)	Tau 640 camera, HD Pro Webcam C920	Mails	63	GLM model	*measured interaction between stress and email condition (stress standardized as z-scores)
[29] (2020)	FLIR Boson 320LW, Intel RealSense D415	Driving	10	SVR with RBF kernel, ROC analize (for 2-level classification)	77%
[13] (2020)	PixFly camera, Specim VNIR, FLIR SC7600	TSST test, Stroop Color-Word test	42	SVM	95.56%
[14] (2019)	FLIR, ECG, OEG-SpO2	Arithmetic mental tasks	10	Random forest, stepwise method, SVM	76.5%
[22] (2017)	FLIR ONE, spectrogram	SCWT test, Mental Computation test	8	CNN	84.59%
[15] (2020)	FLIR SC7600, HSI system	Physical stress test (squats)	20	Linear discriminant, Logistic regression, KNN, decision tree, ensemble learning, SVM	82.11%
[20] (2013)	FLIR, webcam Microsoft	ANU StressDB	45	SVM	72%

Study	Tools	Stressors	Subject	Classifiers/ Algorithms	Best Accuracy Achieved
[18] (2020)	FLIR A310	TSST test	30	ROI comparison	* analysis of the body inflammatory activity and the face peripheral temperature
[30] (2016)	FLIR SC7600, Garmin heart monitor, Miroxi HBR	TSST test	41	DEFP algorithm	*finding parallelism between thermal imprint and stress markers

Source: own elaboration

When analyzing the data from Table 1, it can be noticed that researchers most often use: FLIR cameras, stressor tests such as TSST. The number of respondents is not very large (8-63), although a larger number would allow for a wider spectrum of analyzes and capture the correlation. Invariably, the most frequently used classifier is SVM, the performance of which is often compared by researchers compared to other classifiers, e.g. kNN. Some researchers focus on the comparative analysis of selected ROIs, e.g. of the face or the search for correlation of temperature with other physiological signals. The whole thing makes aware that thermal imaging of stress is developmental and requires further experiments that will, among others: be carried out on a larger number of respondents, use other types of stressors than tests, use newer types of classifiers, e.g. fuzzy logic, use low-cost thermal imaging cameras, especially smartphones.

4. Stress detection – popular solutions

The use of physiological signals such as electromyogram (EMG), electrocardiogram (ECG), skin conductance (GSR), electroencephalogram (EEG) to detect stress is a popular trend among many researchers. Subsequent research works propose improved stress detection algorithms. Unfortunately, many popular methods use high-cost and specialized devices that are unavailable to the common person. Testing is often limited to laboratory conditions, which can affect the realistic nature of the test (compared to home conditions). Table 2 presents a summary of popular methods of stress detection that are an alternative to thermography (selected works taking into account the latest research approaches).

Table 2. List of the most popular methods of stress detection (selected works)

Study	Methods	Description	Classifier
[31] (2020)	EMG, ECG	Signal measurements in 34 subjects, stress induced by mental arithmetic tasks and SCWT test – time pressure and stressful environment. Notice the high efficiency of the EMG signal for binary or multi-level stress detection	Combination of feature selection algorithms (FS) and SVM classification model (with 10 times cross-validation). On two, three, four levels, respectively: 100%, 97.3%, 95.4%
[32] (2019) *[33] (2020)	ECG, EDR	Using a database from the Physionet platform with an implementation in MATLAB Usage of features: RR, QT, EDR *[33] – a similar study: optimized SVM model using a decision tree	SVM – 98.6% accuracy (with Gaussian Kernel Function and all available features)
[34] (2020)	EDA	Monitoring the stress of 41 patients before surgery (electrodermal activity). Determining low, medium and high stress in a local schema	A novel localized supervised learning scheme (the adaptive partitioning of the dataset). Accuracy: 85.06%.
[35] (2020) *[36] (2019) **[37] (2020)	EDA	The use of Samsung GearS smartphones, Empatica E4 smartband, NASA-TLX questionnaire in 32 respondents. Analyzes in EDA Explorer software, MATLAB *[36] – primary stress tests through smartphones and wearable sensors with continuation in [37, 35]	Initial classification: min – 91%, max. 94.52% for 3 classes. After using the Decision-Level Smoothing method: 81.82%, 82.70% for 2 class. Additional logical rules: 94.44% (HR), 100% (EDA)
[38] (2020)	PPG	PPG registration for 14 people performing mental arithmetic (MAT) tasks	Use: the SFFS algorithm, QDA, SVM
[39] (2018) *[40] (2020)	Breath patterns and Kinect	Using Microsoft Kinect to track fluctuations in the deduction of respiratory signals, the skeleton and joints of the human body in 20 subjects performing 3 different tasks: listening to relaxing music, performing exercises, testing the words by Stroop Color-Word test (state equivalents: relaxation, physical stress, mental stress	Use of Fisher's Classifier and the Leave-One-Subject-Out Method. Differentiation of states at the level of: relaxation – 97%, mental stress – 80%, physical stress – 83%. *[40] continuation of the study, 84 subjects, use of random forest and increase in accuracy: 93.90%, 93.40%, 89.05% (states analogy)

Study	Methods	Description	Classifier
[41] (2018) *[42] (2020)	Breath patterns and the bioradar	Reception of respiratory signals by the BioRASCAN-4 bioradar in 43 subjects relaxing in the first stage of the study, and in the second stage – solving mathematical tasks (standard mental stress test) *[42] a proposal for an improved bio-radar method using the Cat Boost classifier (89% accuracy was achieved)	Application of RQA and perceptron three-layer neural network. Achieving an accuracy of 94.4%
[43] (2019)	PCG, ECG	Automatic diagnosis of mental stress in 32 respondents before a stressful exam. Use of the STAI Form Y questionnaire	Use of Kruskal-Wallis statistical test, LS-SVM with 10-fold cross-validation, RBF function. Accuracy: 93.14%
[44] (2020) *[45] (2020)	SOM maps	Proposition of personalized models based on a self-organizing SOM map that can be applied to laboratory and field data. Using WESAD data (laboratory data) and VTT project together with data from smartwatch and smartphone (field data) *[45] use of WESAD and bio-inspired tank computer (RC), consisting of 50 multiplex neurons - 93% accuracy with RCML classifier	After obtaining the SOM results, the use of three groupings: Gaussian Mixture Models (GMM), K-Means, Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN). The best clustering for SOM –K-means. Accuracy (customization various levels): up to 60% (field data), 92% (lab data)
[46] (2019) *[47] (2020)	Twitter	A model that measures attributes from the tweet level (like, comment, forward), and from the user level (publishing behavior) *[47] similar research, use of textual, visual and social attributes	Use of CNN (text-only tweets/retweets), RNN (text and emoticons tweets/retweets)
[48] (2019)	Facial expressions	Detection of negative stress based on emotions: anger, sadness, fear	The use of multitasking cascade convolutional networks – MTCNN
[17] (2018) *[49] (2021)	*iStress*	Stress detection system from: sweating, body temperature, pace of movement. Transferring data from sensors to the IoT cloud *[49] – a similar system based on bioelectronics measuring GSR and skin temperature (SKINTRONICS system)	Mamdani-type fuzzy logic, accuracy: 97%

Source: own elaboration

5. Stress studies – research limitations and gaps

Thermography ensures high registration accuracy, which is partly due to the specific requirements for measuring devices (thermal cameras). These requirements include: no additional emissivity and heat sources, no air drafts in the test room, minimum humidity above 50%, constant room temperature (preferably 21°C). In the case of thermographic stress testing, these requirements cause certain limitations, which nevertheless positively translate into measurement accuracy. The most important limitations are: recommendation to conduct the test in a suitable room, testing outdoors only under appropriate conditions (e.g. avoiding direct sunlight); when using low-budget smartphone thermal imaging cameras – min. 30–50 cm from the examined person and stationary, proper calibration (calibration), sometimes multiple; different temperature measuring range. It is worth bearing in mind the measurement accuracy declared by the manufacturer, which correlates with the measurement error.

When analyzing a number of research works in the field of stress detection, research gaps were noticed, resulting mainly from the diversity of research approaches and psychological and biological aspects of stress. The most important identified research gaps are:

1. Mathematical assumption of stress induction in all respondents in the same way – certainly this assumption facilitates group data analysis, but when referring to human psychology, the assumption becomes wrong because it does not take into account the individual characteristics of a person, his temperament, experiences and psycho-physical condition. Therefore, researchers should collaborate with psychologists as field experts to achieve reliable results and increasingly better stress detection systems.
2. The variety of stress-inducing methods and conditions – many researchers use ready-made stress research tests, generally available, such as TSST, Stroop Color-Word Test. In this way, some of the respondents may have earlier knowledge about the stress tasks they will be performing. And this later translates into the credibility of the results. Another issue is the test conditions – some researchers do not provide information on the conditions under which stress was induced, there is a lack of standardization of the research protocol.
3. Diversity of hypotheses, research devices and classifications – even similar studies are difficult to compare, especially when there are discrepancies in the number of participants, the number of stress classes or validation for a given classification. This can lead to errors in the comparative analysis.
4. Real data sets and ready-made databases – some researchers use publicly available collections instead of creating their own. The ready set is public, but it is characterized by uncertainty as to whether the data was obtained reliably. Added to this is the issue of human diversity in terms of origin, mentality and cultural behavior. Therefore, it is worth building own data sets that take into account human characteristics typical of the environment. The problem here is

the lack of publicity for their own data for other researchers. However, with good will, it can be regulated by direct contact with the creator of the collection.
5. Accuracy of detection systems in real and laboratory conditions – some researchers strive to achieve high accuracy at the same time ignoring the authenticity. The laboratory approach most often ensures high accuracy, but the real conditions, despite the lower accuracy, are more authentic. This is especially important in studies taking into account home or field conditions. Physiological features recorded in controlled laboratory conditions are difficult to transfer to the real world, where people in real life can face various stressors in a more uncontrolled and unusual way.

Certainly stress themes will be developed by the researchers and assisted by modern technology. The development of detection systems based on fuzzy logic, forecasting and quantifying stress, is considered prospective. In addition, it is worth correlating physiological signals with psychometric data to obtain greater reliability of the results. It is also important to consider the research approach that differentiates the level of stress induction, its intensity and differentiation (mental stress, physical stress). Stress detection due to its interdisciplinarity requires a multifaceted approach to the subject.

6. Summary

Ubiquitous health care is an approach that can solve problems of strained health services through the use of preventive strategies. Therefore, many researchers, looking from the perspective, undertake research to support this approach by using low-budget, generally available personal devices, such as a mobile phone. Increased availability of personal monitoring devices and their ease of use can help in a preventive strategy by playing a control, preventive and even diagnostic role. Certainly, prevention requires fewer resources than treatment. At the same time, it responds to the needs of an aging society, higher incidence of diseases due to stressful lifestyle. Thermal imaging can monitor, for example, breathing problems in the elderly, breathing rhythm in children at risk of cot death, irregular breathing in people working under pressure, altered body temperature due to nervousness or stress, discussed in this article. The effectiveness of thermal imaging cameras has been proven in many areas. In the case of thermography of the human body, one should take into account a thoughtful research approach and capture all factors that may deceive the thermal imaging system, e.g. drinking a hot drink before the test, being pregnant (increased body temperature), warming the body with a hot bath or rubbing the skin vigorously. Thermal imaging measures the temperature of the skin, not the internal temperature. Many scientists, studying stress and its role in realistic conditions, recognize the subject of thermal imaging of stress as developmental but difficult to define unequivocally. Nevertheless, there are still innovative solutions related to the use of thermography in the field of stress

detection and in other topics, for example: in ophthalmology [50] (measurement of intraocular pressure – IOP), in everyday life – thermal preferences of residents and thermal comfort [51]. Further development of thermography is expected, using particularly low-cost equipment, available to everyone (home conditions) and not requiring highly specialized knowledge for operation (comprehensibility, ease of use).

Bibliography

[1] Scherz W.D., Baun J., Seepold R., Madrid N.M., Ortega J.A., A portable ECG for recording and flexible development of algorithms and stress detection, Procedia Computer Science, 2020, 176: 2886–2893. DOI: 10.1016/j.procs.2020.09.265.

[2] Shanmugasundaram G., Yazhini S., Hemapratha E., Nithya S., A Comprehensive Review on Stress Detection Techniques, 2019, In: Proceedings of International Conference on System, Computation, Automation and Networking (ICSCAN). DOI: 10.1109/ICSCAN.2019.8878795.

[3] Gedam S., Paul S., Automatic Stress Detection Using Wearable Sensors and Machine Learning: A Review, 2020, 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT). DOI: 10.1109/ICCCNT49239.2020.922569.

[4] Using Drones in Wuhan, https://www.businessinsider.com/china-uses-drones-to-patrol-citizens-during-wuhan-coronavirus-outbreak-2020-1?IR=T

[5] FLIR, camera, https://www.flir.eu. Last accessed: 20.08.2021.

[6] Scanway, https://scanway.pl. Last accessed: 20.08.2021.

[7] Athena Security, https://www.athena-security.com. Last accessed: 20.08.2021.

[8] Vodafone UK. https://newscentre.vodafone.co.uk/press-release/get-the-uk-back-to-work-safely-with-iot-heat-detection-camera/

[9] Panicker S.S., Gayathri P., A survey of machine learning techniques in physiology based mental stress detection systems, 2019. Biocybernetics and Biomedical Engineering, 2019, 39, 444–469. DOI: 10.1016/j/bbe/2019.01.004.

[10] Mizuno T., Sakai T., Kawazura S., Asano H., Akehi K., Itakura N., Measuring facial skin temperature changes caused by mental work-load with infrared thermography, IEEE Transactions on Electronics, Information and Systems, 2016, 136 (11): 1581–1585.

[11] Giannakakis G., Grigoriadis D., Giannakaki K., Review on psychological stress detection using biosignals, 2019. IEEE Transactions on Affective Computing, PP(96): 1–1. DOI: 10.1109/TAFFC.2019.2927337.

[12] Zhang Q., Chen X., Zhan Q., Yang T., Xia S., Respiration-based emotion recognition with deep learning, Comput Ind, 2017, 92, 84–90.

[13] Liu X., Shan Y., Peng M., Chen H., Chen T., Human Stress and StO2: Databases. Features and Classification of Emotional and Physical Stress, 2020. Entropy 2020, 22(9): 962. DOI: 10.3390/e22090962.
[14] Kaga S., Kato S., Extraction of useful features for stress detection using various biosignals doing mental arithmetic, 2019 IEEE 1st Global Conference on Life Sciences and Technologies (LifeTech 2019), 2019, 153–154. DOI: 10.1109/LifeTech.2019.8883967.
[15] Liu X., Xiao X., Cao R., Chen T., Evolution of Facial Tissue Oxygen Saturation and Detection of Human Physical Stress, 2020 Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC), 2020, 144–147. DOI: 10.1109/IPEC49694.2020.9115140.
[16] Rodriquez-Arce J., Lara-Flores L., Portillo-Rodriquez O., Martinez-Mendez R., Towards an anxiety and stress recognition system for academic environments based on physiological features. Computer Methods and Programs in Biomedicine, 2020, 190, 105408. DOI: 10.1016/j.cmpb.2020.105408.
[17] Rachakonda L., Sundaravadivel P., Mohanty S.P., Kougianos E., Ganapathiraju M., A Smart Sensor in the IoMT for Stress Level Detection, 2018 IEEE International Symposium on Smart Electronics Systems (iSES), 2018, 141–145. DOI: 10.1109/iSES.2018.00039.
[18] Cruz-Albarran I.A., Rodriquez-Medina D.A., Leija-Alva G., Dominguez-Trejo B., Osornio-Rios R.A., Morales-Hernandez L.A., Physiological stressor impact on peripheral facial temperature, Il-6 and mean arterial pressure, in young people, Journal of Thermal Biology, 2020, 91. DOI: 10.1016/j.jtherbio.2020.102616.
[19] Gomez de Mariscal E., Munoz-Barrutia A., de Frutos J., Gonzalez-Marcos A.P., Ugena Martinez A. M., Infrared Thermography Processing to Characterize Emotional Stress: A Pilot Study, 2017. VIII International Conference of Pattern Recognition Systems (ICPRS 2017).
[20] Sharma N., Dhall A., Gedeon T., Goecke R., Modeling Stress Using Thermal Facial Patterns: A Spatio-Temporal Approach, 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction, 2013, 387–392. DOI: 10.119/ACII.2013.70.
[21] Zhang H., Feng L., Li N., Jin Z., Cao L., Video-Based Stress Detection through Deep Learning, Sensors 2020, 20, 552. DOI: 10.3390/ s2019552.
[22] Cho Y., Bianchi-Berthouze N., Julier Simon J., DeepBreath: Deep Learning of Breathing Patterns for Automatic Stress Recognition using Low-Cost Thermal Imaging in Unconstrained Settings, 2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII), 2017, 456–463. DOI: 10.1109/ACII.2017.8273639.

[23] Cho Y., Bianchi-Berthouze N., Julier S.J., Marquardt N., Thermsense: Smartphone-Based Breathing Sensing Platform Using Noncontact Low-Cost Thermal Camera, 2017. Available from: arXiv preprint arXiv:1710.05044.
[24] Cho Y., Bianchi-Berthouze N., Oliveira M., Holloway C., Julier S., Nose Heat: Exploring Stress-induced Nasal Thermal Variability through Mobile Thermal Imaging, 2019. 8th International Conference on Affective Computing and Intelligent Interaction(ACII). DOI: 10.1109/ACII.2019.8925453.
[25] Cho Y., Julier SJ., Bianchi-Berthouze N., Instant Stress: Detection of Perceived Mental Stress Through Smartphone Photoplethysmography and Thermal Imaging, 2019. JMIR Mental Health 2019 6(4): e10140. DOI: 10.2196/10140.
[26] Pavlidis I., Tsiamyrtzis P., Shastri D., Wesley A., Zhou Y., Lindner P., Buddharaju P., Joseph R., Mandapati A., Dunkin B., Bass B., Fast by Nature – How Stress Patterns Define Human Experience and Performance in Dexterous Tasks, Scientific Reports, 2012, 2 (305). DOI: 10.1038/ SREP00305.
[27] Bara C.P., Papakostas M., Mihalcea R., A Deep Learning Approach Towards Multimodal Stress Detection, 2020. In Proceedings of the AAAI-20 Workshop on Affective Content Analysis, New York, USA.
[28] Akbar F., Bayraktaroglu A.E., Buddharaju P., Da Cunha Silva D.R., Gao G., Grover T., Gutierrez-Osuna R., Cooper Jones N., Mark G., Pavlidis I., Storer K., Wang Z., Wesley A., Zaman S., Email Makes You Sweat: Examining Email Interruptions and Stress with Thermal Imaging, Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 2019, 668, 1–14. DOI: 10.1145/3290605.33000898.
[29] Cardone D., Perpetuini D., Filippini Ch., Spadolini E., Mancini L., Chiarelli A.M., Merla A., Driver Stress State Evaluation by Means of Thermal Imaging: A Supervised Machine Learning Approach Based on ECG Signal, Applied Sciences, 2020, 10(16): 5673. DOI: 10.3390/ app10165673.
[30] Hong K., Hong S., Real-time stress assessment using thermal imaging, The Visual Computer, 2016, 32, 1369–1377. DOI: 10.1007/s00371-015-1164-1.
[31] Pourmohammadi S., Malaki A., Stress detection using ECG and EMG signals: A comprehensive study, Computer Methods and Programs in Biomedicine, 2020, 193, 105482. DOI: 10.1016/j.cmpb.2020.105482.
[32] Rizwan M. F., Farhad R., Mashuk F., Islam F., Imam M. H., Design of a Biosignal Based Stress Detection System Using Machine Learning Techniques, 2019 International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST), 2019, 364–368. DOI: 10.1109/ICEREST.2019.8644259.

[33] Cruz A.P., Pradeep A., Sivasankar K.R., Krishnaveni K.S., A Decision Tree Optimised SVM Model for Stress Detection using Biosignals, 2020, 2020 International Conference on Communication and Signal Processing (ICCSP). DOI: 10.1109/ICCSP48568.2020.9102043.
[34] Anusha A.S., Sukumaran P., Sarveswaran V., Surees Kumar S.A. Shyam A., Akl Tony J., Preejith S.P., Sivaprakasam M., Electrodermal Activity Based Pre-surgery Stress Detection Using a Wrist Wearable, IEEE Journal of Biomedical and Health Informatics, 2020, 24(1): 92–100. DOI: 10.1109/JBHI.2019.2893222.
[35] Can Y., Chalabianloo N., Ekiz D., Fernandez-Alvarez J., Riva G., Ersoy C., Personal Stress-Level Clustering and Decision Level Smoothing to Enhance the Performance of Ambulatory Stress Detection With Smartwatches, IEEE Access, 2002, 8, 38146–38163. DOI: 10.1109/ ACCESS.2020.2975351.
[36] Can Y.S., Arnrich B., Ersoy C., Stress detection in daily life scenarios using smartphones and wearables sensors: A survey, Journal of Biomedical Informatics, 2019, 92. DOI: 10.1016/j.jbi.2019.103139.
[37] Can Y.S., Gokay D., Reyyan K.D., Ekiz D., Chalabianloo N., Ersoy C., How Laboratory Experiments Can Be Exploited for Monitoring Stress in Wild: A Bridge Between Laboratory and Daily Life, Sensors, 2020, 20, 838. DOI: 10.3390/s20030838.
[38] Zubair M., Yoon Ch., Multilevel mental stress detection using ultra-short pulse rate variability series, Biomedical Signal Processing and Control, 2020, 57. DOI: 10.1016/j.bspc.2019.101736.
[39] Shan Y., Chen T., Yao L., Wu Z., Wen W., Liu G., Remote Detection and Classification of Human Stress Using a Depth Sensing Technique, 2018 First Asian Conference on Affective Computing and Intelligent Interaction (ACII Asia) 2018, 1–6. DOI: 10.1109/ACIIAsia.2018.8470364.
[40] Shan Y., Li S., Chen T., Respiratory signal and human stress: Non-contact detection of stress with a low-cost depth sensing camera, International Journal of Machine Learning and Cybernetics, 2020, 11: 1825–1837. DOI: 10.1007/s13042-020-01074-x.
[41] Machado Fernandez J.R., Anishchenko L., Mental stress detection using bioradar respiratory signals, Biomedical Signal Processing and Control, 2018, 43: 244–249. DOI: 10.1016/j.bspc.2018.03.006.
[42] Anishchenko L., Turetzkaya A., Improved Non-Contact Mental Stress Detection via Bioradar, 2020 International Conference on Biomedical Innovations and Applications (BIA), 2020, 21–24. DOI: 10.1109/BIA50171.2020.9244492.
[43] Cheema A., Singh M., Psychological stress detection using phonocardiography signal: An empirical mode decomposition approach, Biomedical Signal Processing and Control, 2019, 49: 493–505. DOI: 10.1016/j.bspc.2018.12.028.

[44] Tervonen J., Puttonen S.S., Hopsu L., Homorodi Z., Keranen J., Pajukanta J., Tolonen A., Lamsa A., Mantyjaryi J., Personalized mental stress detection with self-organizing map: From laboratory to the field, Computers in Biology and Medicine, 2020, 3124. DOI: 10.1016/ j.compbiomed.2020.103935.
[45] Chandrasekaran S.T., Bhanshali S.P., Banerjee I., Sanyal A., A Bio-Inspired Reservoir Computer for Real-Time Stress Detection From ECG Signal, IEEE Solid-State Circuits Letters, 2020, 3, 290–293. DOI: 10.1109/LSSC.2020.3016924.
[46] Mounika S.N., Kanumuri K.P., Rao K.N., Manne S., Detection of Stress Levels in Students using Social Media Feed, 2019 International Conference on Intelligent Computing and Control Systems (ICCS), 2019, 1178–1183. DOI: 1109/ICCS45141.2019.9065720.
[47] Meshram S., Babu R., Adhikari J., Detecting Psychological Stress using Machine Learning over Social Media Interaction, Proceedings of the 5th International Conference on Communication and Electronics Systems (ICCES 2020), 2020, 646–649. DOI: 10.1109/ICCE48766.2020.9137931.
[48] Zhang J., Mei X., Liu H., Yuan S., Qian T., Negative Emotional Stress Based on Facial Expression in Real Time, 2019 IEEE 4th International Conference on Signal and Image Processing (ICSIP), 2019, 430–434. DOI: 10.1109/SIPROCESS.2019.8868735.
[49] Kim H., Kim Y-S., Mahmood M., Kwon S., Epps F., Rim Y.S., Yeo W-H., Wirelles, continuous monitoring of daily stress and management practice via soft bioelectronics, Biosensors and Bioelectronics, 2021, 173, 112764. DOI: 10.1016/j/bios/2020/112764.
[50] Jędzierowska M., Koprowski R., Wilczyński S., Tarnawska D., The use of infrared thermal imaging in tonometry with a Scheimpflug camera, Journal of Thermal Biology, 2021, 96, 102823.
[51] Li D., Menassa C. C., Kamat V. R., Non-intrusive interpretation of human thermal comfort through analysis of facial infrared thermography, Energy & Buildings, 2018, 176, 246–261. DOI: 10.1016/j.enbuild.2018.07.025.

Szymon Fornal[1], Paweł Karczmarek[2]

Application of the AHP method to analyze the state of knowledge of students in particular years of studying

Abstract: In this study, we thoroughly analyse the state of knowledge of Computer science students at Lublin University of Technology. Analysis was carried out using the well-known Analytic Hierarchy Process (AHP) method of group decision-making. Research focused on subjectively chosen IT technologies in four categories: Object-oriented programming languages, data processing tools, database systems, and Java frameworks. The main idea behind the study is to show the trends in understanding and using the IT technologies by students in dependence of their experience or year of studies. The research was carried using efficient mobile application, also described in this chapter, utilizing the graphical components applied to pairwise comparisons in the AHP process.

Keywords: AHP method, IT industry, university students

1. Introduction

The IT industry is divided into many branches and specializations. Each of them requires adequately prepared programming languages and technologies to work properly. However, the list and form of these technologies is not constant. Over the years, they are systematically changed and improved. Otherwise, they are going to be replaced by newer solutions.

These changes are often noticeable among young IT specialists who are just getting to know the solutions used in the IT industry. For this reason, they tend to use the newest or most up-to-date options. This is partly due to the fact that these technologies are most adapted to the current market realities or the hardware capabilities of computers. Besides, they are constantly supported by their creators.

The group that largely represents young programmers and IT specialists are IT students. For this reason, they were selected for research on the popularity of particular tools and solutions.

The AHP (Analytic Hierarchy Process) method, developed and described by Thomas Saaty [7], can be used to test preferences. It consists in comparing the examined elements in pairs. This way, the subject compares only two things at a time, even if the question checks preferences for more options. The results

[1] Szymon Fornal, Department of Computer Science, Lublin University of Technology, Nadbystrzycka 36B, 20-618 Lublin, Poland; e-mail: s.fornal@op.pl, szymon.fornal@pollub.edu.pl

[2] Paweł Karczmarek, Department of Computer Science, Lublin University of Technology, Nadbystrzycka 36B, 20-618 Lublin, Poland; e-mail: p.karczmarek@pollub.pl

are presented in the form of preference levels. Based on the obtained data, it is necessary to create preference matrices and analyze the obtained results.

Like any AHP method, it also has its drawbacks. They can be corrected by using appropriate solutions, e.g. MAUT (Multi-attribute Utility Theory) [1].

Its biggest advantage is its versatility. It can be applied to many data models, life situations or technologies [8], e.g. in developing facial recognition technology (determining the importance of different parts of the face) [3,6]. Assessing the state of knowledge of students is no exception. Some works approach this from the theoretical and academic side [4]. In this manuscript, focus is put more on analyzing the practical parts of the IT industry, i.e. used languages and technologies.

The results of this research will allow to determine the knowledge of students about the solutions available on the market and will indicate the popularity of these technologies among the youngest group of IT specialists. In addition, the specificity of this type of analysis allows for its development in the future. If the research were continued cyclically at certain (e.g. several years) intervals, it would potentially be possible to determine the trend of changes in the IT market, and thus to predict its future transformations.

The rest of the paper has the following structure. In Section 2, we explain how research data was collected. Section 3 contains results of analysis in numerical and graphical form while Section 4 is intended for conclusions and future work's plans.

2. The method of conducting the study

In order to correctly collect all the data needed for the analysis, it is necessary to create an appropriate environment. As AHP method is based on comparing pairs of attributes with specific preference scales, the best way of collecting data is through sliders, see e.g. [5]. For this reason they were smoothly incorporated into mobile application. Program was created in Android Studio environment using Java language. The data however was stored in NoSQL database, Google Firebase.

Results are more believable and varies if respondents are not aware of rigid scale of comparing preference. To achieve this, sliders were made to be more fluid during use. That is why we used integer values from 0 to 128, despite the fact that preference scale in this study starts with "1:9" and ends on "9:1". Transformation from sliders data to readable data is presented by the following equations.

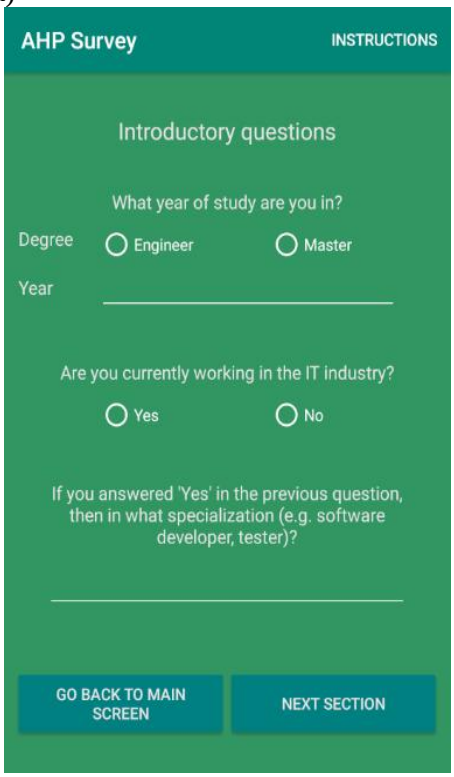
$$y = -\frac{x}{8} + 9, \quad x \in [0, 64), \tag{1}$$

$$y = \frac{x}{8} - 7, \quad x \in [64, 128]. \tag{2}$$

Value *y* is part of result describing user preference, while *x* reflects value taken from slider. For *x* lesser than 64 preference is equal "*y*:1". If *x* is greater than 64, analyzed result has form of "1:*y*" instead. For *x* = 64 preference is set to "1:1" which means no strong feelings towards any presented options.

Examples of application interface are visible on Figures 1(a) and 1(b).

a)



b)

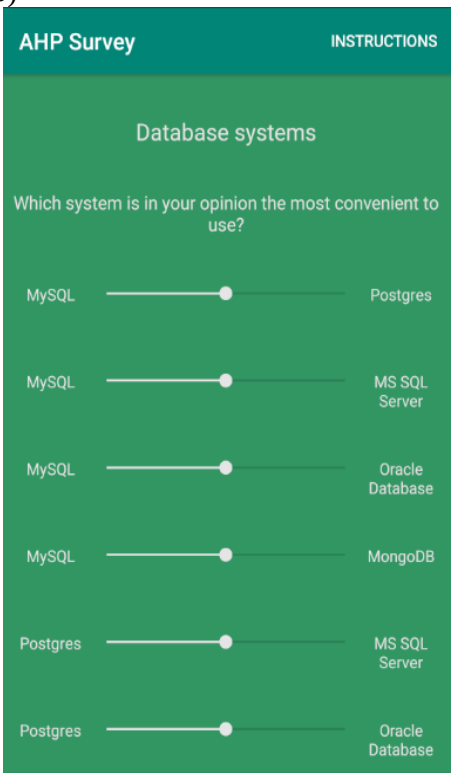


Fig. 1. Interface of application presenting personal data form (a); Interface of application presenting example of question from *Database systems* category (b)

Application included 5 interfaces: one for personal data and 4 for respective categories: object-oriented programming languages, data processing tools, database systems and Java frameworks. Despite the specific order shown in this paper, the categories were randomized in the questionnaires to avoid a situation where any set of questions would have less accurate results due to the fatigue of respondent.

The application was then distributed as an apk executable (which installed application after being downloaded). The questionnaire was distributed to various groups and years of IT students at the Lublin University of Technology.

The research group consisted of 28 students. In terms of the year of study, it included:

- 5 people in the 2nd year of engineering studies;
- 6 people in the 3rd year of engineering studies;
- 8 people in the 1st year of master's studies;
- 9 people in the second year of master's studies.

In terms of employment status:

- 7 people were employed at that time;
- 21 people were not employed during conducting a survey.

On the basis of each respondent's answer, a reciprocal matrix was created, containing user preferences for the pairs assigned to the question. Then, by using the geometric mean, single results were collected in collective reciprocal matrices. This applies to both the general matrix for all users and those for individual groups. Due to the reciprocity property of the matrix, the geometric mean was used to build one output matrix. It is also possible to use the arithmetic mean, but for the results (i.e. for the aggregation of the results obtained from individual experts).

The range of the preference scale is typical for AHP tests and ranges from $1/9$ to 9. Choosing a preference of 9 for a pair of items A and B means an extreme preference for option A. On the other hand, a choice of $1/9$ indicates a complete liking of the option B. A score of 1 means no preferences from the selected options.

To define this scale more precisely, let's denote positive preference as *n* and negative preference as $1/n$. In most studies, values of n in the interval {1, 3, 5, 7, 9} with intermediate values {2, 4, 6, 8} are adopted. Due to the method of collecting the results in this study, this scale is more flexible and ranges between real numbers from 1 to 9 with an accuracy of $1/8$ approximation. This is due to the fact that respondents used high-precision sliders to answer in the application. This method of collecting the results guarantees greater naturalness of the obtained answers (no visible arbitrary scale), as well as increased accuracy of the test results due to the greater precision of the data.

After collecting the results, it is also necessary to define the consistency of the responses of individual experts. When several options are compared, the results may contradict each other. Therefore, it is important to check the so-called inconsistency ratio (in some sources referred to as the consistency ratio). However, to calculate it, it is necessary to know the inconsistency index, which is calculated by the following formula.

$$CI = \frac{\lambda_{max} - n}{n-1}. \tag{3}$$

In this case λ_{max} is maximum eigenvalue of given matrix, while *n* means matrix grade level index.

Then the inconsistency ratio is calculated from formula 4.

$$CR = \frac{CI}{RI}. \tag{4}$$

CI is previously calculated inconsistency index, while *RI* is random consistency index. Its value is taken from the table below. The values in the Table 1 are taken from the article by Gold and Wang [2].

Table 1. The values of random consistency index for each grade of the matrix

Grade of the matrix	**RI**
1	0
2	0
3	0.5799
4	0.8921
5	1.1159
6	1.2358
7	1.3322
8	1.3952
9	1.3952
10	1.4882

The smaller the inconsistency rates, the better the quality of the results. Typically, the results are considered acceptable if their inconsistency ratio is less than 0.1.

3. Results of analysis

Survey was divided into four categories. Each one of them has its unique question set. Full list looks as follows:

1. Object-oriented programming languages.
   a. Which language do you think is more accessible to beginners?
   b. Which language do you think is more universal?
   c. Which language do you think is more convenient to use?
   d. Which language do you think runs faster (application/program speed)?
   e. Which language do you think is more popular in the current state of IT industry?
2. Data processing tools.
   a. Which tool do you think is more convenient for the graphical interpretation of the results?
   b. Which tool do you think is better for mathematical/statistical analysis?
   c. Which tool do you think offers more features?
   d. Which tool do you think is more efficient (processes and analyzes data faster)?

e. Which tool do you think is more popular in the current state of IT industry (only in the data analysis branch)?

3. Database systems.
   a. Which system do you find more convenient to use?
   b. Which system do you think is faster in data management, e.g. in an application?
   c. Which system do you think works better for large projects?
   d. Which system do you think works better for small projects?
   e. Which system do you think offers more features/tools?
   f. Which system do you think is more popular in the current state of IT industry?
4. Java frameworks.
   a. Which framework do you think is more recognizable?
   b. Which framework do you think is better for communication with the database?
   c. Which framework do you think is more complex?
   d. Which framework do you think is lighter (slows down the application less)?
   e. Which framework do you think is better suited for developing applications in the MVC model?

In this paper only one question for each category is being thoroughly analyzed. For the rest of the results, we give only the preference values gathered in the Table 2. Questions about popularity of given technology were chosen because they are more compact and give the best view of summarized respondents' opinions.

Table 2. The preference values for all answers in respective questions

Question number	**Answer**	**Preference (weight)**
1a	Java	0.25
	Python	0.33
	C#	0.19
	C++	0.23
1b	Java	0.38
	Python	0.26
	C#	0.21
	C++	0.15
1c	Java	0.34
	Python	0.3
	C#	0.23
	C++	0.14
1d	Java	0.19
	Python	0.22
	C#	0.23
	C++	0.37

Question number	Answer	Preference (weight)
1e	Java	0.5
	Python	0.26
	C#	0.16
	C++	0.08
2a	Python	0.19
	R	0.21
	MATLAB	0.21
	Excel	0.39
2b	Python	0.17
	R	0.31
	MATLAB	0.29
	Excel	0.23
2c	Python	0.45
	R	0.19
	MATLAB	0.23
	Excel	0.13
2d	Python	0.39
	R	0.27
	MATLAB	0.15
	Excel	0.19
2e	Python	0.52
	R	0.17
	MATLAB	0.13
	Excel	0.19
3a	MySQL	0.34
	PostgreSQL	0.18
	MsSQL	0.2
	Oracle Database	0.15
	MongoDB	0.14
3b	MySQL	0.27
	PostgreSQL	0.18
	MsSQL	0.24
	Oracle Database	0.18
	MongoDB	0.13
3c	MySQL	0.18
	PostgreSQL	0.17
	MsSQL	0.25
	Oracle Database	0.28
	MongoDB	0.11

Question number	Answer	Preference (weight)
3d	MySQL	0.33
	PostgreSQL	0.24
	MsSQL	0.15
	Oracle Database	0.1
	MongoDB	0.18
3e	MySQL	0.16
	PostgreSQL	0.15
	MsSQL	0.28
	Oracle Database	0.29
	MongoDB	0.12
3f	MySQL	0.2
	PostgreSQL	0.18
	MsSQL	0.23
	Oracle Database	0.27
	MongoDB	0.12
4a	Spring	0.48
	Hibernate	0.25
	Play	0.09
	Grails	0.09
	Vaadin	0.09
4b	Spring	0.51
	Hibernate	0.49
4c	Spring	0.46
	Play	0.17
	Grails	0.18
	Vaadin	0.18
4d	Spring	0.34
	Play	0.23
	Grails	0.21
	Vaadin	0.22
4e	Spring	0.58
	Play	0.2
	Grails	0.22

3.1. Object-oriented programming languages

In this paragraph we analyze popularity of programming languages. The following technologies were taken into consideration: Java, Python, C# and C++. Each respondent had to compare every pair of languages (for this question there were 6 pairs in total) and decide which of presented options is in their opinion more popular. In this manner it was possible to collect answers in form of reciprocal matrix. Single matrices were later transformed into one by using geometric mean. Based on this data we were able to calculate each language's

weight. Thanks to this, analyzed data can be presented in more graphical form, visible on Figure 2.

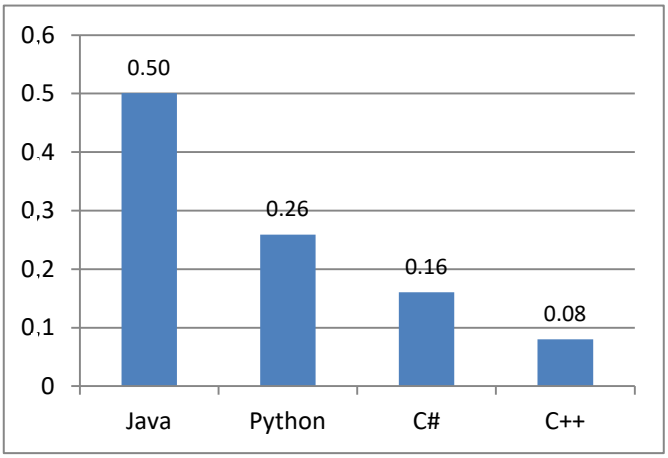


Fig. 2. Chart presenting each language's weight in general group for question "Which language do you think is more popular in the current state of IT industry?"

Source: own elaboration

Results are also visible in the Table 3.

Table 3. The weights and scales of preference for programming languages in general group

	Java	**Python**	**C#**	**C++**	**Weight**
Java	1	2.1	3.6	4.94	0.5
Python	0.48	1	1.55	3.69	0.26
C#	0.28	0.64	1	2.24	0.16
C++	0.2	0.27	0.45	1	0.08

Data can be also presented for more specific groups. In this research it was divided based on student's current degree and on their employment status. Obtained results can be seen on Figure 3, Figure 4 and Tables 4–7.

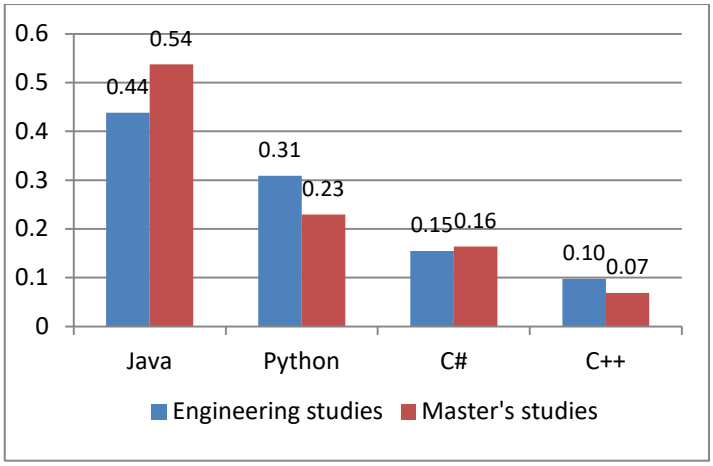


Fig. 3. Chart presenting each language's weight based on current degree for question "Which language do you think is more popular in the current state of IT industry?"

Source: own elaboration

Table 4. The weights and scales of preference for programming languages among engineering degree students

	Java	**Python**	**C#**	**C++**	**Weight**
Java	1	1.08	4.11	4.11	0.44
Python	0.92	1	1.25	3.77	0.31
C#	0.24	0.8	1	1.32	0.15
C++	0.24	0.26	0.76	1	0.1

Table 5. The weights and scales of preference for programming languages among master's degree students

	Java	**Python**	**C#**	**C++**	**Weight**
Java	1	3.21	3.31	5.57	0.54
Python	0.31	1	1.79	3.64	0.23
C#	0.3	0.56	1	3.16	0.16
C++	0.18	0.27	0.32	1	0.07

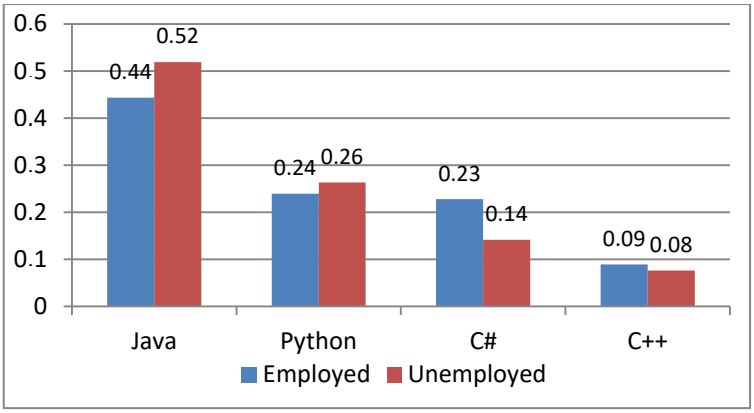


Fig. 4. Chart presenting each language's weight based on employment status for question "Which language do you think is more popular in the current state of IT industry?"

Source: own elaboration

Table 6. The weights and scales of preference for programming languages among employed students

	Java	**Python**	**C#**	**C++**	**Weight**
Java	1	2.58	1.62	4.14	0.44
Python	0.39	1	1.21	3.18	0.24
C#	0.62	0.83	1	2.52	0.23
C++	0.24	0.31	0.4	1	0.09

Table 7. The weights and scales of preference for programming languages among unemployed students

	Java	**Python**	**C#**	**C++**	**Weight**
Java	1	1.96	4.7	5.24	0.52
Python	0.51	1	1.69	3.88	0.26
C#	0.21	0.59	1	2.16	0.14
C++	0.19	0.26	0.46	1	0.08

To check how reliable are these results it is necessary to calculate inconsistency ratio. Its values for respective groups are as follows:

- 0.011 for all students;
- 0.048 for students of engineering studies;
- 0.034 for students of master's studies;
- 0.021 for working students;
- 0.017 for non-working students.

As we stated earlier, the inconsistency is acceptable as long at *CR* value is lower than 0.1. In this scenario all values were situated safely below this threshold, so received data is authoritative.

Based on these results some assumptions can be made. Java was decisively chosen as the most popular object-oriented programming language. This state of affairs may be partly due to the fact that Java is the most frequently used language in the course of IT studies at the Lublin University of Technology. Therefore, most students are familiar with this technology. It is also worth noting some dependencies for individual groups. Undergraduate students very often chose Python as the best option. It is a language that has grown in popularity significantly in recent years. For this reason, many young IT professionals may be inclined to use it. As for the employed, they often mentioned C#, which would mean that this language is very popular on the labor market in Lublin.

3.2. Data processing tools

In this category survey fillers were presented with four options: Python, R, MATLAB, and Excel. The way of collecting and analyzing data is analogous to these presented in previous subsection. Charts with particular tools' weights are presented on Figures 5–7 and in Tables 8–12.

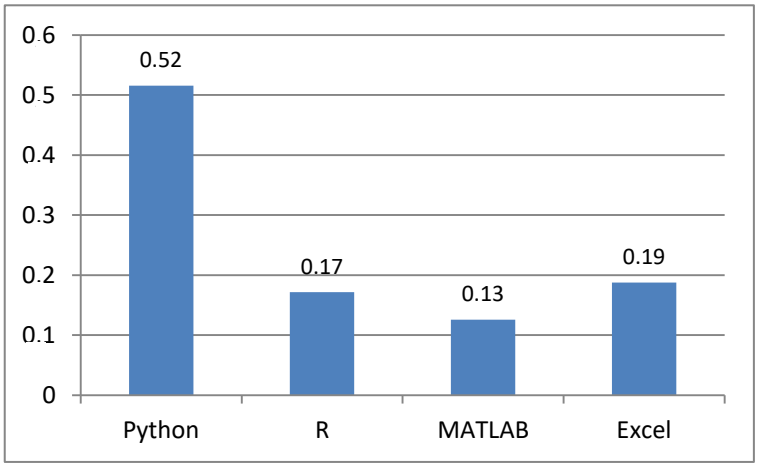


Fig. 5. Chart presenting each data analysis tool's weight in general group for question "Which tool do you think is more popular in the current state of IT industry (only in the data analysis branch)?"

Source: own elaboration

Table 8. The weights and scales of preference for data analysis tools in general group

	Python	**R**	**MATLAB**	**Excel**	**Weight**
Python	1	3.02	4	2.81	0.52
R	0.33	1	1.39	0.9	0.17
MATLAB	0.25	0.72	1	0.67	0.13
Excel	0.36	1.11	1.49	1	0.19

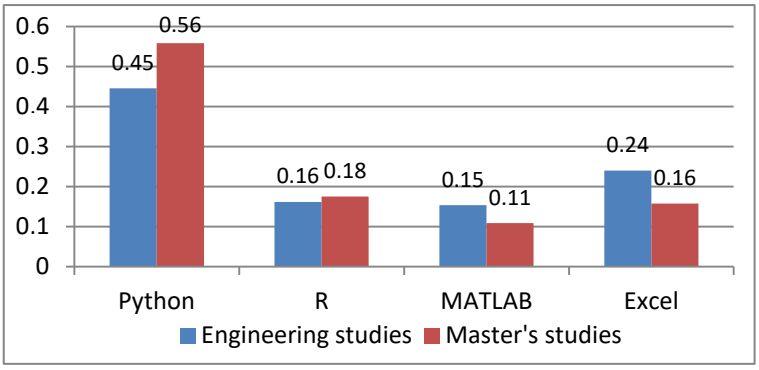


Fig. 6. Chart presenting each data analysis tool's weight based on current degree for question "Which tool do you think is more popular in the current state of IT industry (only in the data analysis branch)?"

Source: own elaboration

Table 9. The weights and scales of preference for data analysis tools among engineering degree students

	Python	**R**	**MATLAB**	**Excel**	**Weight**
Python	1	2.96	3.37	1.49	0.45
R	0.34	1	0.92	0.83	0.16
MATLAB	0.3	1.09	1	0.65	0.15
Excel	0.67	1.2	1.54	1	0.24

Table 10. The weights and scales of preference for data analysis tools among master's degree students

	Python	**R**	**MATLAB**	**Excel**	**Weight**
Python	1	3.06	4.47	4.24	0.56
R	0.33	1	1.82	0.94	0.18
MATLAB	0.22	0.55	1	0.68	0.11
Excel	0.24	1.06	1.47	1	0.16

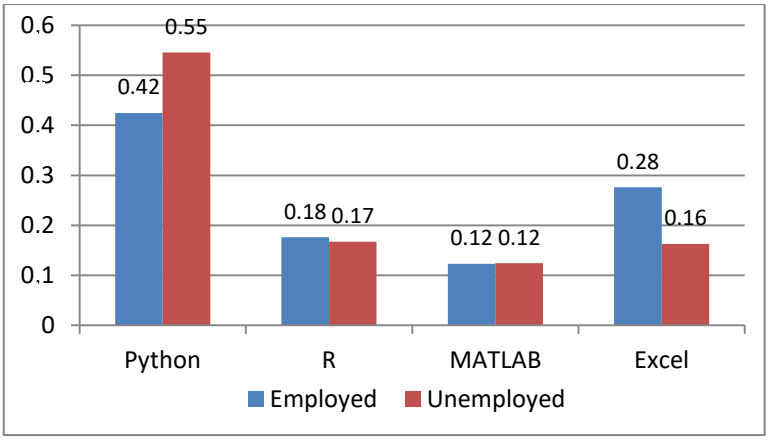


Fig. 7. Chart presenting each data analysis tool's weight based on employment status for question "Which tool do you think is more popular in the current state of IT industry (only in the data analysis branch)?"

Source: own elaboration

Table 11. The weights and scales of preference for data analysis tools among employed students

	Python	**R**	**MATLAB**	**Excel**	**Weight**
Python	1	1.9	2.85	2.25	0.42
R	0.53	1	1.5	0.49	0.18
MATLAB	0.35	0.67	1	0.41	0.12
Excel	0.44	2.04	2.44	1	0.28

Table 12. The weights and scales of preference for data analysis tools among unemployed students

	Python	**R**	**MATLAB**	**Excel**	**Weight**
Python	1	3.52	4.48	3.03	0.55
R	0.28	1	1.36	1.1	0.17
MATLAB	0.22	0.74	1	0.79	0.12
Excel	0.33	0.91	1.27	1	0.16

As before we calculated values of inconsistency ratio. They are as follows:

- 0.0002 for all students;
- 0.013 for students of engineering studies;
- 0.009 for students of master's studies;
- 0.03 for working students;
- 0.002 for non-working students.

This time inconsistency was even lower. It means that respondents were answering questions judiciously and decisively.

Python was considered the most popular among all groups. Such a state of affairs may be due to a fact that Python is also well-known programming language (what was shown in previous category). In this case some respondents may have chosen it based on its overall popularity, not necessarily as data analysis tool. With a clearly worse result, second place was occupied by Microsoft Excel. It was quite popular choice among employed and undergraduate students.

3.3. Database systems

As previously, a few technologies were presented to surveyed. This time answer set contains five options: MySQL, PostgreSQL, MsSQL, Oracle Database, and MongoDB. Results of answers' analysis is presented on below Figures (numbers 8, 9 and 10) and Tables (13–17).

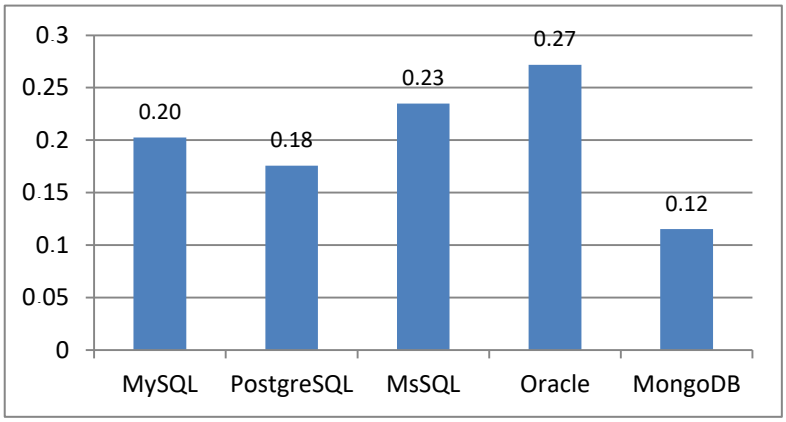


Fig. 8. Chart presenting each database system's weight in general group for question "Which system do you think is more popular in the current state of IT industry?"

Source: own elaboration

Table 13. The weights and scales of preference for database systems in general group

	MySQL	**PostgreSQL**	**MsSQL**	**Oracle**	**MongoDB**	**Weight**
MySQL	1	1.23	0.87	0.53	2.3	0.2
PostgreSQL	0.81	1	0.84	0.54	1.84	0.18
MsSQL	1.15	1.19	1	1.07	1.96	0.23
Oracle	1.89	1.85	0.93	1	1.6	0.27
MongoDB	0.43	0.54	0.51	0.62	1	0.12

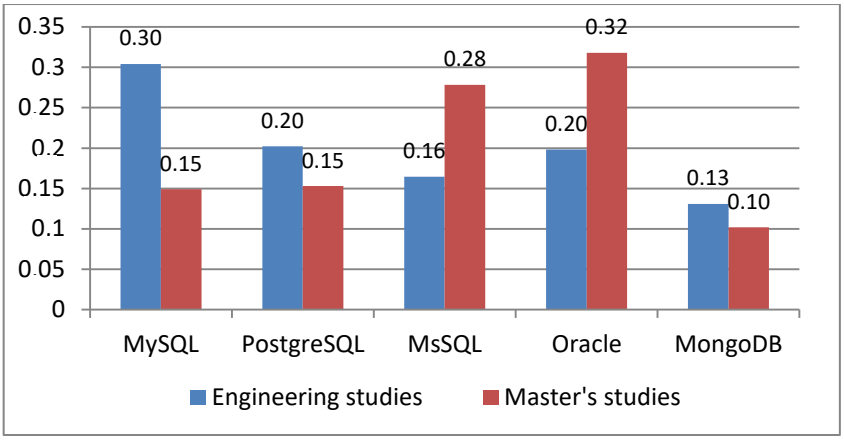


Fig. 9. Chart presenting each database system's weight based on current degree for question "Which system do you think is more popular in the current state of IT industry?"

Source: own elaboration

Table 14. The weights and scales of preference for database systems among engineering degree students

	MySQL	**PostgreSQL**	**MsSQL**	**Oracle**	**MongoDB**	**Weight**
MySQL	1	1.99	1.86	1.11	2.33	0.3
PostgreSQL	0.5	1	1.33	1.1	1.74	0.2
MsSQL	0.54	0.75	1	0.96	1.23	0.16
Oracle	0.9	0.91	1.04	1	1.32	0.2
MongoDB	0.43	0.57	0.81	0.76	1	0.13

Table 15. The weights and scales of preference for database systems among master's degree students

	MySQL	**PostgreSQL**	**MsSQL**	**Oracle**	**MongoDB**	**Weight**
MySQL	1	0.9	0.53	0.33	2.29	0.15
PostgreSQL	1.11	1	0.62	0.34	1.9	0.15
MsSQL	1.88	1.62	1	1.15	2.66	0.28
Oracle	3.07	2.93	0.87	1	1.81	0.32
MongoDB	0.44	0.53	0.38	0.55	1	0.1

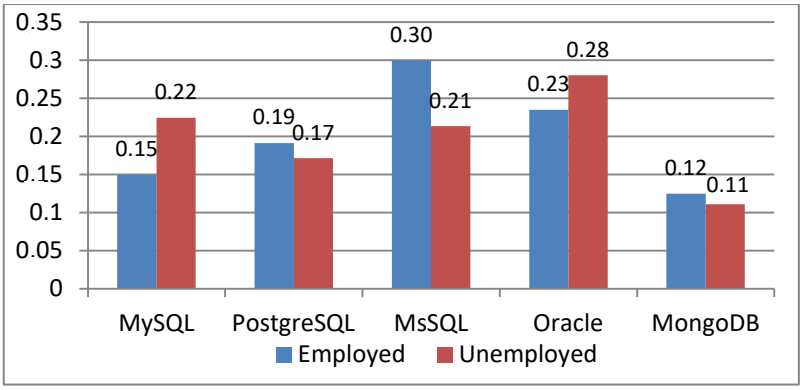


Fig. 10. Chart presenting each database system's weight based on employment status for question "Which system do you think is more popular in the current state of IT industry?"

Source: own elaboration

Table 16. The weights and scales of preference for database systems among employed students

	MySQL	**PostgreSQL**	**MsSQL**	**Oracle**	**MongoDB**	**Weight**
MySQL	1	0.47	0.69	0.44	1.82	0.15
PostgreSQL	2.11	1	0.59	0.64	1.28	0.19
MsSQL	1.45	1.69	1	1.63	2.7	0.3
Oracle	2.29	1.55	0.61	1	1.31	0.23
MongoDB	0.55	0.78	0.37	0.77	1	0.12

Table 17. The weights and scales of preference for database systems among unemployed students

	MySQL	**PostgreSQL**	**MsSQL**	**Oracle**	**MongoDB**	**Weight**
MySQL	1	1.69	0.94	0.56	2.49	0.22
PostgreSQL	0.59	1	0.94	0.51	2.08	0.17
MsSQL	1.07	1.06	1	0.93	1.77	0.21
Oracle	1.78	1.97	1.07	1	1.72	0.28
MongoDB	0.4	0.48	0.57	0.58	1	0.11

Similar to the previous part of the research, the values determining the inconsistency of the answers were calculated. They are presented below:

- 0.021 for all students;
- 0.011 for students of engineering studies;
- 0.041 for students of master's studies;
- 0.044 for working students;
- 0.027 for non-working students.

Results are closer to those presented in first category. Still they are good enough to consider collected data as reliable.

The results for most of the systems were similar. This differs from the previous parts of the research, where the dominance of one option was noticeable. Oracle Database was considered the most popular, and MongoDB – the least. Groups of more experienced students were more inclined towards solutions of Microsoft and Oracle, but engineering students considered MySQL to be more popular. The trend turned out to be more even in groups categorized by the employment status. Here the results are closer to those for the general group. It is worth noting that among the unemployed, MsSQL was the most preferred option.

The high popularity of MsSQL and Oracle may result from their large presence on the market, as well as from the fact that they appear in later years of the Lublin University of Technology study program. Younger programmers are better acquainted with the MySQL system, which is often perceived as technology for beginners. The low popularity of MongoDB can be explained by the fact that it is a NoSQL system. Solutions of this type are very different from other technologies, so many people prefer not to utilize them, because they can use proven solutions instead. It is possible that this trend will change in the upcoming years.

3.4. Java frameworks

Similar as for database systems, respondents had to compare five technologies: Spring, Hibernate, Play, Grails and Vaadin. Their preferences scores were collected and analyzed. Results of this actions are presented on below Figures (11–13) and Tables (18–22).

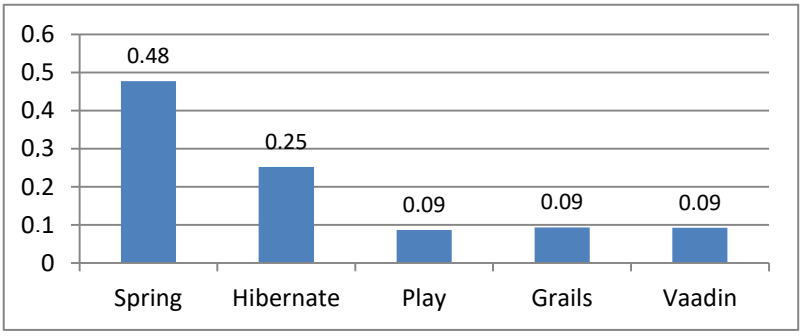


Fig. 11. Chart presenting each Java framework's weight in general group for question "Which framework do you think is more recognizable?"

Source: own elaboration

Table 18. The weights and scales of preference for Java frameworks in general group

	Spring	Hibernate	Play	Grails	Vaadin	Weight
Spring	1	2.84	4.6	4.55	4.35	0.48
Hibernate	0.35	1	3.28	3.29	2.87	0.25
Play	0.22	0.31	1	0.88	0.94	0.09
Grails	0.22	0.3	1.13	1	1.09	0.09
Vaadin	0.23	0.35	1.06	0.92	1	0.09

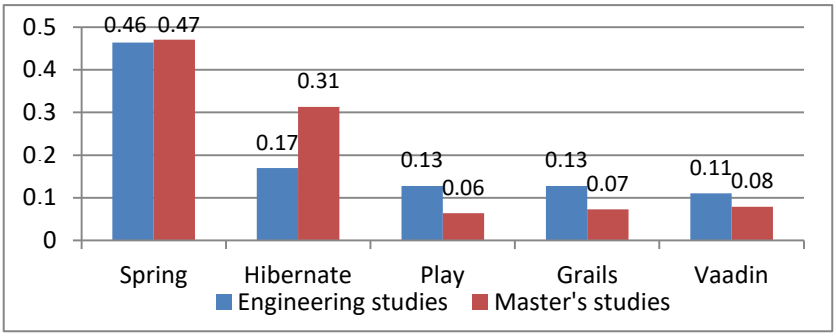


Fig. 12. Chart presenting each Java framework's weight based on current degree for question "Which framework do you think is more recognizable?"

Source: own elaboration

Table 19. The weights and scales of preference for Java frameworks among engineering degree students

	Spring	Hibernate	Play	Grails	Vaadin	Weight
Spring	1	3.34	3.61	3.58	3.47	0.46
Hibernate	0.3	1	1.45	1.54	1.46	0.17
Play	0.28	0.69	1	1	1.29	0.13
Grails	0.28	0.65	1	1	1.33	0.13
Vaadin	0.29	0.68	0.78	0.75	1	0.11

Table 20. The weights and scales of preference for Java frameworks among master's degree students

	Spring	Hibernate	Play	Grails	Vaadin	Weight
Spring	1	2.56	5.39	5.31	5.04	0.47
Hibernate	0.39	1	5.55	5.39	4.44	0.31
Play	0.19	0.18	1	0.81	0.77	0.06
Grails	0.19	0.19	1.23	1	0.96	0.07
Vaadin	0.2	0.23	1.3	1.05	1	0.08

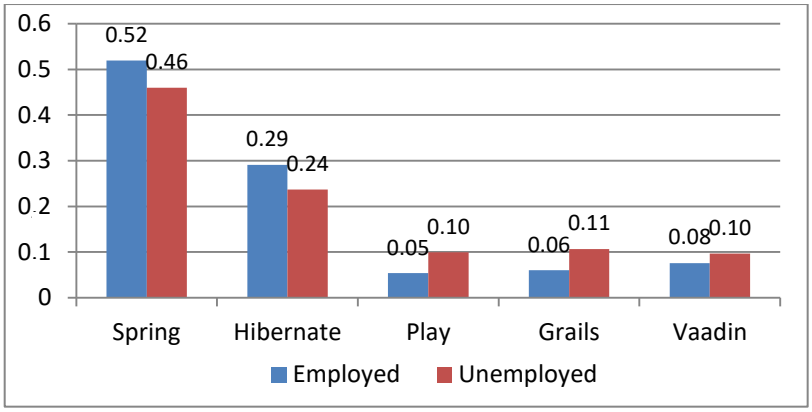


Fig. 13. Chart presenting each Java framework's weight based on employment status for question "Which framework do you think is more recognizable?"

Source: own elaboration

Table 21. The weights and scales of preference for Java frameworks among employed students

	Spring	**Hibernate**	**Play**	**Grails**	**Vaadin**	**Weight**
Spring	1	3.35	6.4	6.37	6.1	0.52
Hibernate	0.3	1	6.28	6.33	4.46	0.29
Play	0.16	0.16	1	0.76	0.67	0.05
Grails	0.16	0.16	1.32	1	0.69	0.06
Vaadin	0.16	0.22	1.49	1.45	1	0.08

Table 22. The weights and scales of preference for Java frameworks among unemployed students

	Spring	**Hibernate**	**Play**	**Grails**	**Vaadin**	**Weight**
Spring	1	2.69	4.13	4.07	3.89	0.46
Hibernate	0.37	1	2.64	2.65	2.48	0.24
Play	0.24	0.38	1	0.93	1.06	0.1
Grails	0.25	0.38	1.08	1	1.27	0.11
Vaadin	0.26	0.4	0.94	0.79	1	0.1

In the last category, same as before it was necessary to calculate inconsistency ratio. Its values equal:

- 0.014 for all students;
- 0.006 for students of engineering studies;
- 0.024 for students of master's studies;
- 0.038 for working students;
- 0.01 for non-working students.

CR stays always below 0.1, so this data is credible.

In this part, the greatest inconsistencies and discrepancies in the results were expected, because despite the high popularity of Java (which is confirmed by the research in the first category), many students may not have had contact with tools for this particular language. Fortunately, the inconsistency of the answers was avoided, which means that the results can be considered reliable.

Undoubtedly, Spring was recognized as the best known framework in all research groups. Hibernate was also highly rated. The rest of the options achieved a negligible score. Cause of this results is that despite the fact that other frameworks are gaining popularity in the IT industry and are used by specialists, students most often deal with Spring and Hibernate, which are even taught during their studies. Additionally, Spring is very complex and can even be divided into smaller modules, e.g. Spring MVC. This stands in opposition to more specialized frameworks focused on their niches. For this reason, Spring is more recognizable, even to those people who do not normally program in Java.

4. Conclusions and future work

The AHP method is very flexible and transparent to use. The research related to checking the knowledge of students was fully possible to plan and implement, and the form of filling in the questionnaires was clear to the respondents. The only drawback regarding the convenience of solving the questionnaire concerns the number of required pairs. Due to the specificity of this method, each analyzed element must be compared with all the others. This means a relatively greater amount of work for the respondent than in the case of traditional surveys. For example, if 4 technologies were compared, the number of sliders that had to be moved was 6. In the case of 5 elements, there were already 10 sliders. Their number increases proportionally according to the formula, i.e.

$$n(n-1)/2\,, \tag{5}$$

where *n* is number of items compared.

A hand-made mobile application was used to conduct this study. This format allows greater flexibility in setting the sensitivity of the sliders. This has two benefits. First, when moving the slider, the respondent is unaware that he or she is moving on a rigid scale from $1/9$ to 9. Thanks to that, the results are more natural and less forced by the interface. The second advantage is the fact that the preference scores are also more precise, which makes them more reliable and less prone to an approximation error. The fact of using own application also gives some control over the collected results (they are not stored on an external site, which can hold them even after withdrawing the consent of the person conducting the test).

The downside to using the application is how it is transmitted. In order to be able to complete the survey, it is first necessary to install it via an apk file. This extension is reserved for Android executables. Modern phones by default block downloading such files for security reasons (if they are not installed by, for example, the Play Store). In this case, downloading the application may be cumbersome for respondents, and many of them may not even trust the program. This situation is particularly difficult during a pandemic, when there is no direct contact with the person conducting the study, and thus it is not known whether the file comes from a trusted source. Additionally, unlike web applications, mobile applications are dedicated to specific systems and it is impossible to collect data from e.g. owners of Apple smartphones. Unfortunately, these issues translated into a satisfactory, but still not very large research group.

Due to their specificity of comparing elements in pairs, AHP tests may lead to some inconsistencies in the answers. If the respondent sets the following preferences for items A, B and C (for the sake of example, the exact preference scales have been omitted): A > B, B > C and C > A, we are talking about inconsistent results. In this case, the data may be illogical or misleading. To avoid such hassles, an inconsistency ratio is used. Its value determines whether the given results are consistent and valuable for the analysis. It is generally accepted that the safe *CR* value should be less than 0.1. This coefficient was also calculated in the conducted research. Fortunately, in all groups the rates were low enough (after aggregating the expert responses), which means that the groups of respondents responded consistently and no results needed to be rejected.

The results for individual categories were discussed in the course of chapters. To summarize, the most popular technologies in their categories are: Java (object-oriented programming languages), Python (data analysis tools) and Spring (Java frameworks). Database systems yielded definitely more even results (except for the less popular MongoDB). As the study was conducted on students, great impact on the results were subject to their contact with technology as part of the course of study. Of course we have to consider that in this article only popularity surveys are presented in detailed way. When other factors are taken into account, the results clearly differ.

The results of these studies reflect the current state of affairs and are dependent on the actual trends in the IT market. In the coming years, it would be worth repeating the analysis and observing how the preferences of IT students change over a longer period of time.

Overall, the AHP method is very flexible. It can be used in many cases, using scales with different levels of precision. There are also no problems with analyzing data for different groups. In the case of this study, the groups were divided according to the degree of study or employment status, but it would also be possible to use the criteria of age, gender, origin or work experience.

The AHP research itself is transparent for the respondents, although more time-consuming than ordinary surveys.

The future work directions are, among others, an application of fuzzy versions of AHP or extending the research to the wider group of students, not only from the Lublin University of Technology.

Bibliography

[1] Dyer J.S., Remarks on the Analytic Hierarchy Process, Managements Science, 1990, 36(3): 249–258.

[2] Golden B.L., Wang Q., An Alternative Measure of Consistency, B.L. Golden, A. Wasil & P.T. Harker (eds.) Analytic Hierarchy Process: Applications and Studies, 1990, 68–81.

[3] Karczmarek P., Kiersztyn A., Pedrycz W., An Application of Graphic Tools and Analytic Hierarchy Process to the Description of Biometric Features, ICAISC 2018, LNAI 10842, 2018, 137–147.

[4] Karczmarek P., Pedrycz W., Czerwiński D., Kiersztyn A., The Assessment of Importance of Selected Issues of Software Engineering, IT Project Management, and Programming Paradigms Based on Graphical AHP and Fuzzy C-Means, 2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), 2020, 1–7.

[5] Karczmarek P., Pedrycz W., Kiersztyn A., Fuzzy Analytic Hierarchy Process in Graphical Approach, Group Decision and Negotiation, 2021, 30(2): 463–481

[6] Karczmarek P., Pedrycz W., Kiersztyn A., Rutka P., A Study in Facial Features Saliency in Face Recognition: An Analytic Hierarchy Process Approach, Soft Computing, 2017, 21, 7503–7517.

[7] Saaty T.L., How to Make a Decision: The Analytic Hierarchy Process, Europan Journal of Operational Research 1990, 48, 9–26.

[8] Saaty T.L., Vargas L.G., Models, Methods, Concepts & Applications of the Analytic Hierarchy Process Second Edition, 2012.

Adrian Jaczyński[1], Paweł Karczmarek[2]

Analysis of the use of e-learning platforms in connection with the COVID-19 virus

Abstract: In this study, we focus mainly on the analysis of the use of e-learning platforms during the COVID-19 pandemic. The research is based on the analysis of the popularity of individual e-learning platforms based on publicly available data from Google Trends website. Additionally conducted was a research survey focusing on the impact of the pandemic on the quality of education. There were considered the most common and popular e-learning platforms. We present the results of research in the community of students, mainly of computer science faculties. The analysis of the answers can shed the light on the advantages and shortcomings of the e-learning process and platforms.

Keywords: e-learning platforms, COVID-19, remote learning

1. Introduction

In November 2019, a new type of coronavirus called SARS-CoV-2 was detected. In many cases, it causes acute respiratory infectious disease COVID-19. A few months after its detection, the virus turned into a global pandemic that affects almost all countries in the world. A significant number of countries, wishing to contain the pandemic, introduced countermeasures, which led to the closure of their economies and traditional school services. As one can see, the COVID-19 pandemic has forced all educational institutions such as universities, high schools, technology, elementary schools, and even kindergartens to switch to remote learning. All institutions are forced to introduce various online teaching methods such as e-learning systems or mobile learning applications. Therefore, it is so important to use tools that allow us to study the dynamics of the pandemic, carry out as many systematic screening tests as possible to check whether a person is or has been sick, and collect the data accurately.

Online learning is new to almost all learners as well as teachers. Conducting classes in this form is a very big challenge for teachers who have to learn how to use selected programs, as well as transfer knowledge without the possibility of direct help for students, as was the case with face-to-face contacts. In the case of younger primary school audiences, parents must also play a large role in their teaching process, which often leads to many conflicts on the parent-teacher line.

[1] Adrian Jaczyński, Department of Computer Science, Lublin University of Technology, Nadbystrzycka 36B, 20-618 Lublin, Poland; e-mail: adrian.jaczynski@pollub.edu.pl
[2] Paweł Karczmarek, Department of Computer Science, Lublin University of Technology, Nadbystrzycka 36B, 20-618 Lublin, Poland; e-mail: p.karczmarek@pollub.pl

To mitigate the effects of online learning, various educational interventions have been undertaken using available online tools and even traditional media such as television.

Available e-learning tools play one of the most important roles in online education during the pandemic. They help plan, manage, deliver and track the entire learning process. They also facilitate communication between students and teachers, so one can easily share all kinds of materials needed to conduct classes. A very big advantage of most systems is that they are simply free, so educational institutions and the students themselves do not have to worry about all kinds of subscriptions needed to use the software to carry out classes. Students can use such systems via web browsers or by installing appropriate applications on their own device. The only requirement is the need to have an internet network to be able to take full advantage of the application's capabilities. For this purpose, for example, cellular networks or local wireless networks can be used.

Providing study materials has become one of the main challenges for many universities during the pandemic. Initially, many universities tested selected e-learning platforms to optimally match the possibilities offered by a given platform to the form of teaching. Most systems allow you to share your own screen or video calls with the use of webcams, which largely allows you to check the independence of students and whether they are involved in the process of conducting classes.

The main goal of this study is to analyze the popularity of e-learning platforms during the pandemic. For this purpose, it will be analyzed which e-learning platform has increased its popularity the most due to the pandemic and which is the most popular. An analysis will also be made of the periods in which the platforms gained significant popularity growth and how the pandemic affected student performance.

The rest of the manuscript is as follows. In the second section discussed is the general idea of e-learning platforms. Section 3 describes Google Trends platform and the program to download data. Section 4 presents the test results and analysis of the collected data. The last section presents the conclusions and plans for future work.

2. E-learning platform description

Today, there is an increasing demand for online learning technologies such as e-learning. E-learning is a group effort of people such as teachers, administrators, designers and users from various fields of science, who share their knowledge together, allowing people to learn from their experience [1, 10]. This chapter describes different e-learning applications that can create virtual learning spaces.

Microsoft Teams is the digital cloud application center. The application provides the possibility of conducting conversations, video meetings and file sharing in one system designed to facilitate online learning or business meetings at work [14]. Often multiple people must work with the same documents or be able to access the same information at the same time. As a rule, work is spread over many locations, which often causes many problems. The solution offered by Microsoft offers a lot of time-saving and efficiency-increasing features that can be used by organizations of all sizes, both companies and educational institutions. The program can be treated as one large "super application" that integrates many different applications into one, while offering a user-friendly interface [12].

Cisco Webex is web video conferencing software. The application was developed by Cisco and allows users to create virtual meetings and join them immediately without having to schedule them. The software provides support for video calls that can be joined by up to 200 users. Due to the fact that the application is mainly used by corporations, it offers many security functions, such as data or message encryption [16]. Cisco runs the application in its own infrastructure, so all you need to access the application is a browser and an internet connection. The solution is delivered as SaaS (Software as a service) software, which makes it easy to scale to any number of websites and projects. Thanks to this, the application can be adapted to the needs of every company, regardless of its size, by immediately scaling for a significant number of users [5].

Zoom application is a cloud service that offers features such as online meetings, messaging and file transfer, and secure session recording. Like other platforms, it offers the possibility of real-time communication with people from different parts of the world via a computer or a mobile device. The key advantage of the application is the ability to securely record or store sessions without using third-party software [3]. Unlike the Skype application, Zoom does not require participants to have an account or download the program to their computer, all one need to do is link to the meeting to join it via a web browser. The application also includes password protection to ensure confidentiality and recording capabilities. In addition, the software allows you to record the meeting in two files: audio and video. Thanks to this solution, you can easily share the recording due to a much smaller file size. In this way, the privacy of users who do not want their face to be visible in the recorded meeting can be protected [11].

Moodle is the brainchild of Martin Dougiamas, who designed the application while working on his Ph.D. thesis which dealt with a socio-constructivist approach to learning. In this way, Moodle originally distinguished itself with features such as a discussion forum or a very user-friendly interface [6]. Currently, the application allows teachers to share tasks, documents, quizzes, workshops, chat and the entire forum for students in a way that ensures high-quality learning. It is one of the most flexible applications among the free solutions available worldwide. This system has been specifically designed to

help teachers create high-quality online courses to impart knowledge [2]. As the application is available in the open source model, it is developed by many specialists from around the world. It contains all information about the development of the application and a guide to the source code, thanks to which every new developer can easily find himself in the application and contribute to its development. The application can be downloaded in various packages, which ensures its flexibility and the possibility of selecting the appropriate version for a given institution [7, 20].

Skype has been downloaded over 200 million times worldwide by approximately 85 million users. The user base grows by over 100,000 a day, and from 3.5 to 4 million users which use the application at one time. The application provides real-time multimedia services via the Internet [13]. This solution is ideal for individual and group conversations and can be used from a variety of devices such as computers, tablets and mobile phones. The software offers support for messages, voice calls and video in HD quality [17]. The application also enables encryption of connections and messages to prevent unauthorized users from accessing confidential information. Skype uses a proprietary session establishment protocol which is designed to verify the user's identity and enable peering to agree on a session secret key. Thanks to this, people communicating with each other use the key to obtain confidential communication during the session [4].

Universality is a platform connecting the academic environment with the business environment, founded in 2017 by a Polish startup. The very idea for the application was born in 2014, when Jerzy Czepiel founded his first IT company, conducting classes at the Jagiellonian University. This tool was designed to connect employers, academic teachers and students in one ecosystem, while increasing the level of education [19].

Udemy is a web-based course sharing application. Each course includes lectures, which may be videos, slides or text. Content instructors can add various resources and practice exercises to serve as teaching aids. The application is available via a web browser as well as on portable and mobile devices. All available courses are available on request, so the course can be started at any time and the user has unlimited time to complete it. If the user does not like the course, the platform allows for a refund within 30 days from the date of purchase [18].

Google Meet is a platform that provides a service for making video calls. Anyone with a Google account can create a virtual meeting that lasts up to 60 minutes and allows a user to invite up to 100 users. Companies, universities, schools, and other organizations can use specific advanced features such as the ability to organize meetings for more people (up to 250 participants) and the ability to conduct live broadcasts for up to 100,000 viewers in a given domain [9].

3. Data preparation

3.1. Google Trends

Google Trends is a web application from Google commonly used to research the popularity of keywords in a given period. On the website, one can learn about the prevailing trends regarding individual topics that are searched by Internet users in the Google search engine. The application allows to check whether a keyword is searched more or less often by users. Data is collected from all over the world and becomes useful for website owners or entrepreneurs who want to deliver content and products to users in their area of interest. The values presented by Google Trends reflect the number of searches for a given term in relation to the total number of all searches in the Google search engine.

3.2. Test application

Data from the Google Trends website were used to analyze the popularity of individual e-learning platforms. For this purpose, a data retrieval application was created in Python. In order to be able to download data in the program, the Pytrends library was used – an unofficial Google Trends API, which allows to extract useful data downloaded from the website with a simple interface [15]. At the beginning, all passwords for which data were to be retrieved were read from the *.csv file in the application. Then, for each single password, a function from the Pytrend library was called, getting data. The individual parameters used in the build_payload function mean:

- kw_list – a list of passwords for which data is to be downloaded. At this point, we can enter a maximum of 5 passwords;
- timeframe – a period for which data are to be downloaded;
- geo – a location for which data should be downloaded. By default, they are searched for the whole world, but you can also narrow down the data download to e.g. one country. The application saves the downloaded results to a *.csv file.

4. Test results and analysis

4.1. Survey results

The survey was mainly addressed to students and learners. 130 people replied to the questionnaire. It was mainly attended by students from the Lublin University of Technology who were at different levels of studies. For the multiple-choice question regarding the use of platforms at universities, the following results were obtained (Figure 1):

- Microsoft Teams – 128 respondents (98.5%);

- Moodle – 90 people surveyed (69.2%);
- Zoom – 11 people surveyed (8.5%);
- Webex – 2 respondents (1.5%);
- Google Meet – 1 respondent (0.8%).

The survey shows that the largest number of respondents uses the Microsoft Teams and Moodle platforms. This is due to the fact that the respondents are mainly students of the Lublin University of Technology.

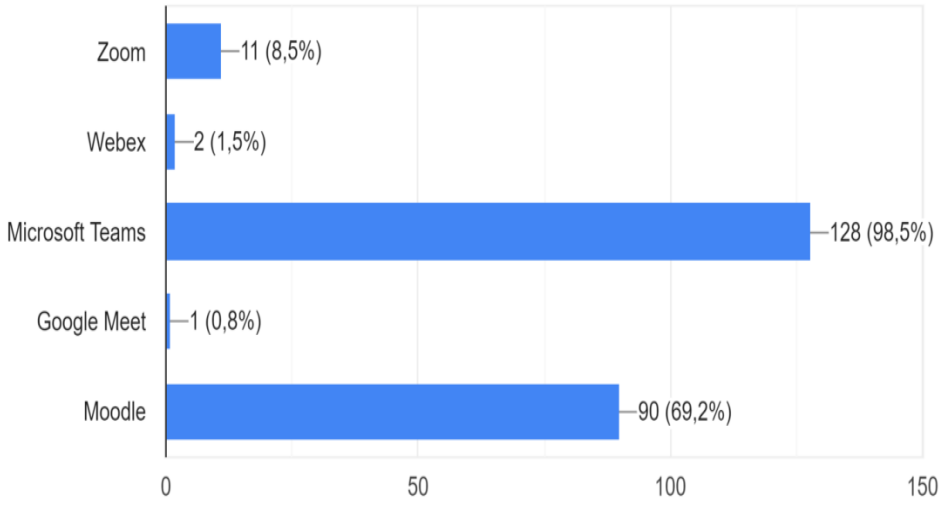


Fig. 1. The most used platform

Source: own elaboration

The next question was about assessing one's own academic achievement, both before and during the pandemic, as shown in Figure 2. As can be seen, in the course of the pandemic, the majority of respondents stated that their academic performance is definitely better or just better.

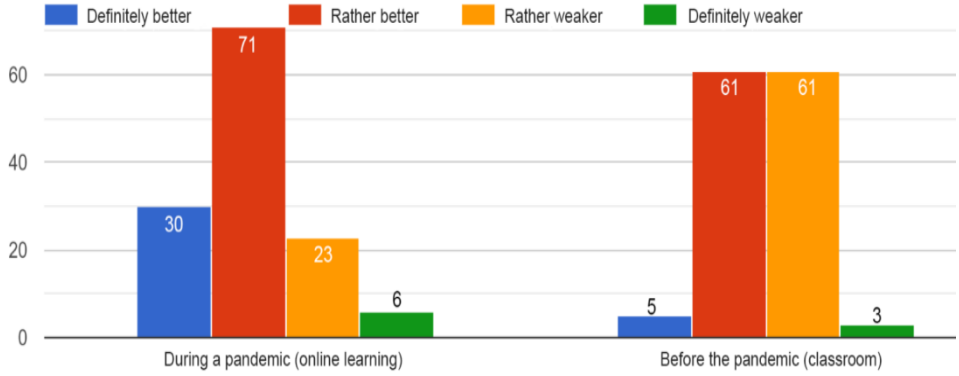


Fig. 2. Assess academic performance before and during a pandemic

Source: own study

Results for the question "Do you think it is easier to get better results before or during a pandemic?" are as follows (Figure 3):

- definitely during a pandemic – 38 people surveyed (29.2%);
- rather during a pandemic – 49 respondents (37.7%);
- rather before the pandemic – 12 respondents (9.2%);
- definitely before the pandemic – 2 respondents (1.6%);
- I have no opinion – 29 respondents (22.3%).

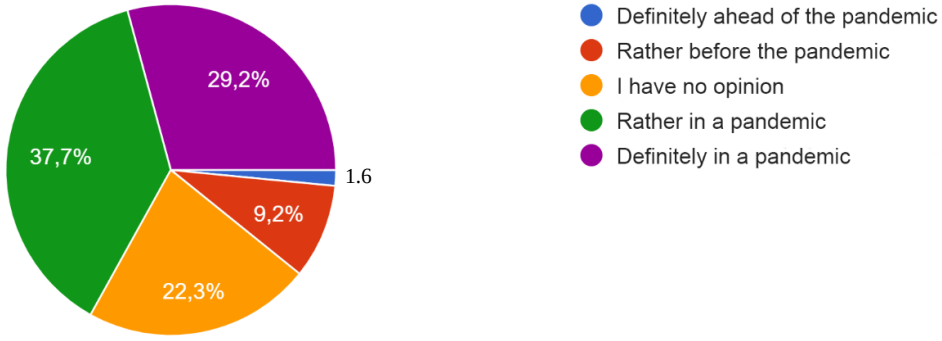


Fig. 3. Assessment of achieving better academic results

Source: own study

Table 1 shows the percentage results for the frequency of use of individual teaching aids before the pandemic. The results regarding the use of teaching aids during a pandemic are presented in Table 2.

Table 1. Use of scientific aids before the pandemic

	Very often	**Often**	**Rarely**	**Very rarely**	**Never**
Books	3.8%	10%	30%	40%	16.2%
Tutorials	36.9%	44.6%	15.4%	3.1%	–
Thematic blogs	16.2%	46.2%	25.4%	8.5%	3.8%
Finding solutions on the Internet (Stackoverflow)	65.4%	30%	3.8%	–	0.8%
Consultation with the teacher	3.1%	16.9%	46.9%	20.8%	12.3%
E-learning platforms (Udemy etc.)	7.7%	16.9%	27.7%	20%	27.7%

Table 2. Use of scientific aids during a pandemic

	Very often	**Often**	**Rarely**	**Very rarely**	**Never**
Books	4.6%	8.5%	23.8%	34.6%	28.5%
Tutorials	52.3%	31.5%	13.8%	2.3%	–
Thematic blogs	26.2%	45.4%	16.2%	6.9%	5.4%
Finding solutions on the Internet (Stackoverflow)	79.2%	16.9%	2.3%	0.8%	0.8%
Consultation with the teacher	4.6%	17.7%	32.3%	31.5%	13.8%
E-learning platforms (Udemy etc.)	12.3%	23.1%	19.2%	18.5%	26.9%

Tables 3 and 4 present the responses of the respondents regarding the assessment of forms of contact with the teachers, both before and during the COVID-19 epidemic.

Table 3. Assessment of forms of contact with teachers before the pandemic

	Definitely positive	**Rather positive**	**Rather negative**	**Definitely negative**	**None**
Stationary meetings	32.3%	60%	6.2%	–	1.5%
Telephone	2.3%	9.2%	9.2%	6.2%	73.1%
Teleconferences	3.8%	10%	6.2%	3.1%	76.9%
Videoconferencing (Skype etc.)	4.6%	14.6%	5.4%	1.5%	73.8%

Table 4. Assessment of forms of contact with teachers during the pandemic

	Definitely positive	Rather positive	Rather negative	Definitely negative	None
Stationary meetings	8.5%	10%	4.6%	3.1%	73.8%
Telephone	3.1%	14.6%	6.9%	6.9%	68.5%
Teleconferences	13.1%	45.4%	3.8%	2.3%	35.4%
Videoconferencing (Skype etc.)	25.4%	59.2%	5.4%	0.8%	9.2%

The assessment of particular aspects of distance learning during the pandemic is presented in Table 5. A very large proportion of the respondents stated that during distance learning it is possible to gain new competences or learn about new technologies related to distance learning. The biggest problems that can be noticed during remote learning are problems with the Internet connection, computer hardware, or even stress related to the lack of time, which was assessed mostly negatively by the respondents.

Table 5. Assessment of various aspects of the transition to remote learning

	Definitely positive	Rather positive	Rather negative	Definitely negative	No opinion
Opportunity to gain new competences	18.5%	43.1%	21.5%	6.2%	10.7%
The opportunity to learn about new technologies	16.2%	47.7%	23.1%	6.9%	6.2%
Internet connection problems	5.4%	10%	30.8%	33.1%	20.8%
Problems with computer hardware	7.7%	13.1%	32.3%	24.6%	22.3%
Stress related to lack of time	10.8%	13.1%	20%	35.4%	18.5%

Table 6 presents the results concerning the possibilities of maintaining balance between private life and science. In the case of resignation from a hobby for science, 60 respondents answered that they agreed with this statement, while the rest stated that they do not have to give up their favourite activities during the pandemic. A significant part of the respondents also stated that in order to learn well, they must devote their professional duties, which may be caused by the lack of an adequate amount of time. Lack of integration or social meetings after the end of the classes has a negative impact on the motivation to learn, which about 75 respondents agree. People who have to travel to the university agreed that during the pandemic they could save on renting a flat.

Table 6. Possibility of maintaining a balance between private life and study

	Definetly agree	Tend to agree	Rather disagree	Strongly disagree	No opinion
Giving up hobby/passion for science	9.2%	36.9%	25.4%	25.4%	3.1%
Learning takes place at the expense of professional responsibilities	12.3%	28.5%	16.9%	16.9%	25.4%
No integration meetings	26.2%	31.5%	12.3%	16.9%	13.1%
No socializing after class	30.8%	27.7%	16.2%	16.2%	9.2%
Saving on renting premises closer to the university	30%	16.2%	10.8%	10.8%	32.3%

To the question "How do you evaluate the contact with other students during group work?" the respondents answered as follows (Figure 4):

- definitely positive – 51 respondents (39.2%);
- rather positive – 59 respondents (45.4%);
- rather negative – 18 people surveyed (13.8%);
- definitely negative – 2 respondents (1.5%).

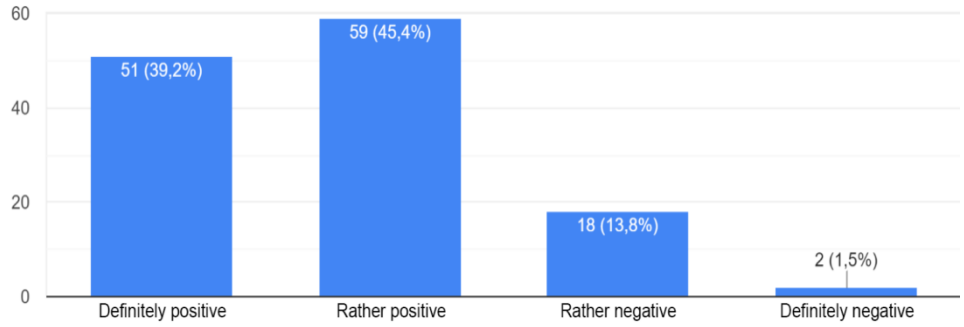


Fig. 4. Assessment of contact with other students during group work

Source: own study

Figure 5 presents a graph of the assessment of university preparedness for the transition to remote mode. The respondents answered as follows:

- definitely positive – 31 respondents (23.8%);
- rather positive – 68 respondents (52.3%);
- rather negative – 24 respondents (18.5%);
- definitely negative – 7 respondents (5.4%).

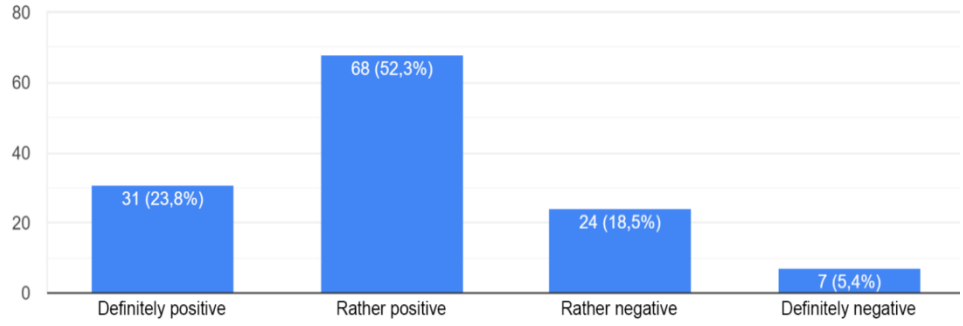


Fig. 5. Assessment of the university's preparation for the transition to distance learning

Source: own study

Figure 6 shows the results of the survey on how respondents perceive the lack of face-to-face contact with the leaders. They are as follows:

- definitely an advantage – 14 respondents (10.8%);
- more like an advantage – 33 respondents (25.4%);
- rather a disadvantage – 48 respondents (36.9%);
- definitely a disadvantage – 35 respondents (26.9%).

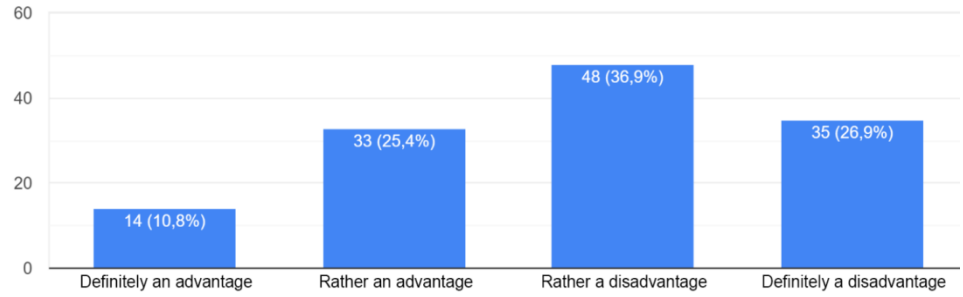


Fig. 6. Assessment of the lack of face-to-face contacts with the lecturers

Source: own study

4.2. E-learning platform popularity research results

This section will analyze the popularity of e-learning platforms. The data was downloaded from the Google Trends website and the popularity scale for each checked password is the same. The value of 100 means the highest popularity of the password. The value 50 means that the popularity of the password is twice lower, while the value 0 indicates that there is not enough data for the given password [8].

Figure 7 shows a graph of the popularity of the *Microsoft Teams* search term. As can be seen in 2019, even before the COVID-19 pandemic, the popularity was around 10. This means that very few people around the world searched Google for any information about a given e-learning platform. In 2020, the situation is completely different. At the turn of March/April, during the lockdown, the distance learning system was introduced in almost all educational institutions. As can be seen in the chart, it was during this period that the popularity of searching for a given term reached the value of 100. This means that at that moment a lot of people from all over the world started searching for information about a given e-learning platform. During the summer holidays, one may notice a double drop in popularity, which is caused by a break from studying. In September, however, another increase in the popularity of search can be noted, which is caused by returning to school. During this period, popularity was around 65 and began to decline again closer to Christmas and the winter break.

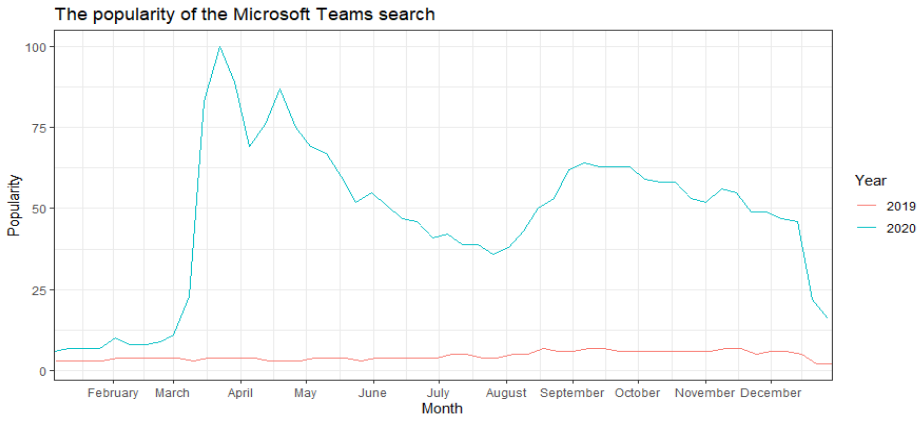


Fig. 7. The popularity of the Microsoft Teams search

Source: own study

A graph showing the popularity of the *Webex* platform is shown in Figure 8. The popularity of this e-learning platform in 2019 was around 15. It can be seen that it is slightly more compared to Microsoft Teams. In this case, this platform also achieved the greatest popularity at the turn of April/May 2020, when educational institutions began to switch to distance learning due to the COVID-19 pandemic. Since then, the popularity of the Webex platform has dropped steadily to around 45 during the holiday season. After the holidays, one can observe that the popularity of the platform has not increased drastically, which

may indicate that most users, such as universities, simply did not apply the offer that the platform offered. This could be influenced by many factors, such as optimization, services offered, the security of users' personal data or even the ability to use the application via a web browser. The second significant increase in popularity can only be observed at the turn of November/December, just before the Christmas break, and it amounted to around 60.

The popularity of the search term for the *Zoom* e-learning platform is illustrated in Figure 9. As can be seen in 2019, the popularity was around 10. This is due to the low interest in using e-learning platforms. In 2020, this platform was most popular at the turn of April. In the following months, until the end of the holiday period, the popularity of the Zoom platform was systematically decreasing, to the level of around 30. After the holidays, a slight increase in popularity was noted to the level of around 50. Over the next months, until the holiday season, the popularity of the platform began to decline again, only to stop at around 30 before Christmas.

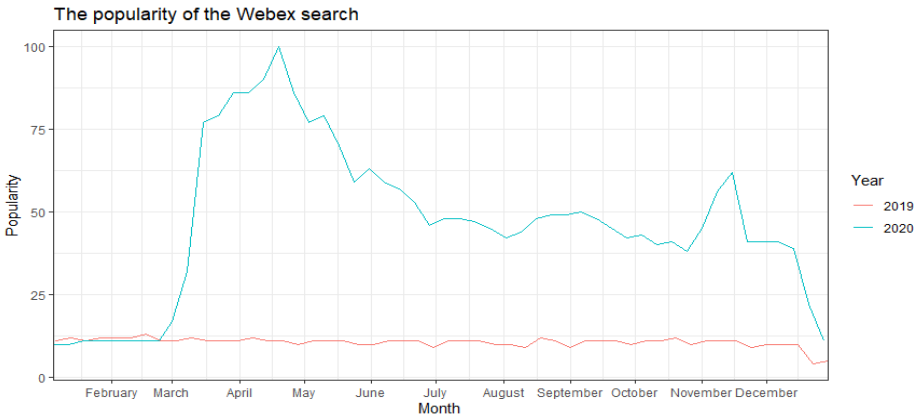


Fig. 8. The popularity of the Webex search

Source: own study

Figure 10 shows the popularity of the Skype platform. In 2019, the popularity of the platform was around 20. Before the COVID-19 pandemic in 2020, the popularity was around 10 and it saw a significant increase in popularity to 100 in March/April during the outbreak of the pandemic. Thereafter, the popularity of the Skype platform began to drop drastically to around 20 during the holiday season. After the summer holidays, the popularity did not increase anymore and remained at the same level as in 2019, which may mean that most universities started using other solutions (Microsoft Teams, Zoom or Webex).

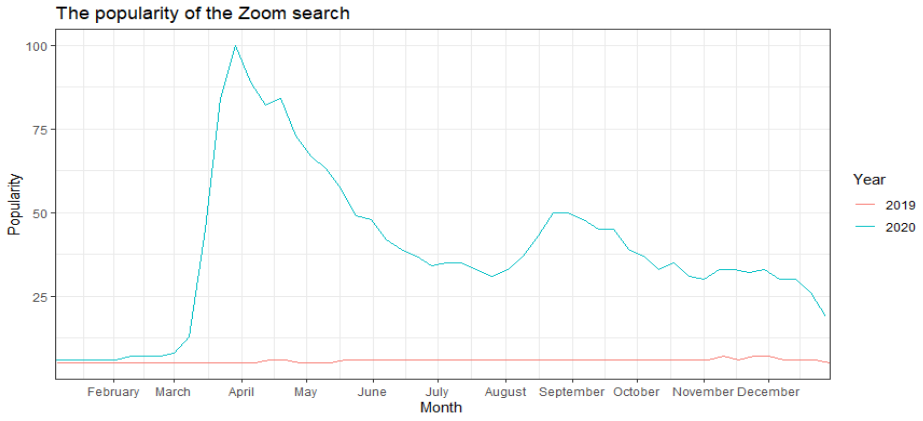


Fig. 9. The popularity of the Zoom search

Source: own study

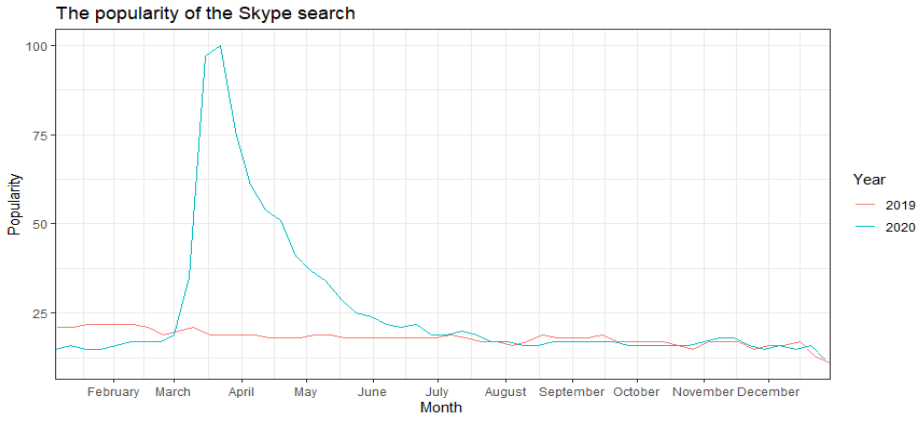


Fig. 10. The popularity of the Skype search

Source: own study

4. Conclusions

The conducted analysis allowed to answer the following research question 1: "Which e-learning platform increased its popularity most in connection with the pandemic?". According to the popularity charts, it can be said that the Microsoft Teams platform has increased its popularity the most. This is due to the fact that the popularity of the platform in 2020, after the holiday season, remained at the

level of around 65. No other platform has achieved such popularity result. In addition, this is due to the fact that in 2019 the popularity of the platform in the same period was only around 10.

Verifying Research Question 2: "Which e-learning platform is the most popular during the COVID-19 pandemic?" it can be said that the most popular platform is Microsoft Teams. This is evidenced by the results of the survey (Figure 1), which show that as many as 128 (98.5%) of the respondents answered that this platform is most often used at their universities. Additionally, when comparing the Microsoft Team search results with other platforms, it can be stated that it is the most used platform.

The analysis of the popularity of individual e-learning platforms gave an answer to research question 3: "In which periods did e-learning platforms have the greatest increase in popularity?". When analyzing the popularity charts of various e-learning platforms, it can be seen that they recorded the greatest increase at the turn of March and May 2020. During this period the popularity of the platforms reached 100. This is due to the introduction of distance learning in the context of the COVID-19 pandemic. All educational institutions, adapting to the situation, were forced to check various types of offers of individual platforms, which clearly contributed to achieving such a popularity result.

According to the respondents (Figure 3), better results in science could be achieved during the pandemic. The respondents replied that it was easier to achieve better results in education rather during a pandemic (37.7%) and definitely during a pandemic (29.2%). Also, a significant proportion of respondents noticed an improvement in their own learning achievements during the pandemic, as shown by the responses (Figure 2), where 71 respondents (54.6%) stated that during the pandemic they achieved better learning outcomes than before the pandemic. Additionally, 30 respondents (23.1%) stated that they achieve much better results in science. Therefore, the analysis of the survey gave an unambiguous answer to research question 4: "How did the pandemic affect the results achieved by students?".

Based on the responses of the respondents, it can be concluded that remote learning provides the opportunity to acquire new competences and the opportunity to learn new technologies. Students do not have to travel to universities, so they have more time to develop their own hobbies and can save more money if they do not have to rent an apartment. The disadvantages of distance learning are frequent problems with the Internet connection, computer equipment and stress related to the lack of free time (Table 5). During the pandemic, there are no integration meetings and it negatively affects the well-being of students (Table 6).

As a future work we are going to continue the research on more numbers of people, not only from the Lublin area. Moreover, we are going to compare the effectiveness of the e-learning platforms.

Bibliography

[1] Aijuan D., Honglin L., Ontology-Based Information Integration in Virtual Learning Environment, International Conference on Web Intelligence, 2005, 762–765.

[2] Al-Ajlan A., Zedan H., Why Moodle?, 12th IEEE International Workshop on Future Trends of Distributed Computing Systems, 2008, 58–64.

[3] Archibald M.M., Ambagtsheer R.C., Casey M.G., Lawless M., Using Zoom Videoconferencing for Qualitative Data Collection: Perceptions and Experiences of Researchers and Participants. International Journal of Qualitative Methods, 2019, 18: 1–8.

[4] Berson T., Skype security evaluation, Anagram Laboratories, 2005.

[5] Cisco WebEx LLC: Why WebEx? Discover the benefits of real-time web collaboration in your organization, http://cache42.prolifiq.com/core/f63fefee-1a55-4692-9587-9df80066aaca/Why%20WebEx.pdf. Last accessed: 22.08.2021.

[6] Dougiamas M., Moodle: A Virtual Learning Environment for the Rest of Us, TESL-EJ, 2004.

[7] Dougiamas M., Moodle, www.moodle.org. Last accessed: 23.08.2021.

[8] Google Trends, https://trends.google.pl/. Last accessed: 23.08.2021.

[9] Google Meet, https://apps.google.com/intl/pl/meet/. Last accessed: 23.08.2021.

[10] Graham S., Simeonov S., Boubez T., Davis D., Daniels G., Nakamura Y., Neyama R., Building Web Services with Java: Making Sense of XML, SOAP, WSDL and UDDI, 2001.

[11] Gray L.M., Wong-Wylie G., Rempel G.R., Cook K., Expanding Qualitative Research Interviewing Strategies: Zoom Video Communications, The Qualitative Report, 2020, 1292–1301.

[12] Hubbard M., Bailey M. J., Mastering Microsoft Teams. End User Guide to Practical Usage, Collaboration, and Governance, 2018.

[13] Kuan-Ta C., Chun-Ying H., Polly H., Chin-Laung L., Quantifying Skype User Satisfaction, ACM SIGCOMM Computer Communication Review, 2006, 399–410.

[14] Martin L., Tapp D., Teaching with Teams: An Introduction to Teaching an Undergraduate Law Module Using Microsoft Teams, Innovative Practice in Higher Education, 2019.

[15] Pytrends, https://github.com/GeneralMills/pytrends. Last accessed: 23.08.2021.

[16] Singh R., Awasthi S., Updated Comparative Analysis on Video Conferencing Platforms Zoom, Google Meet, Microsoft Teams, WebEx Teams and GoToMeetings, EasyChair, 2020.

[17] Skype, https://www.skype.com/pl/about/. Last accessed: 23.08.2021.

[18] Udemy, https://support.udemy.com/hc/pl/articles/229232187-Jakdzia%C5%82a-Udemy-Cz%C4%99sto-zadawane-pytania. Last accessed: 23.08.2021.

[19] Universality, https://universality.io/o-nas/#wiecej. Last accessed: 23.08.2021.

[20] Williams B., Dougiamas M., Moodle for Teachers, Trainers and Administrators, http://www.moodle.org/, 2005.

Adam Kiersztyn[1], Krystyna Kiersztyn[2], Patrycja Jędrzejewska-Rzezak[3], Paweł Karczmarek[4], Witold Pedrycz[5]

Analysis of self-awareness of beginning programmers

Abstract. One of the key problems in creating study programmers in IT studies is the choice of a programming language to teach the basics of programming and algorithmic. In this paper we would like to look at this problem from the point of view of preferences of the most interested, i.e. the students starting their studies in Computer Science. The main aim of the research is to examine the preferences of the students and to verify the research hypothesis that students entering education are aware of their decisions and know the needs of the labor market, and are therefore able to identify the appropriate languages to learn based on both their experience as well as other people's opinions and general trends. The research uses a questionnaire, but instead of a simple question about the best programming language the Analytic Hierarchy Process method is used, in which comparisons are made in pairs.

Keywords: Analytic Hierarchy Process (AHP), decision-making, first programming language, computer science education.

1. Introduction

The problem of choosing the right language to teach the basics of programming to students starting IT studies is still an open issue [3], [6], [8], [12]–[16], [19]–[22]. Many students starting their studies have some basic knowledge of programming [4], [7], a significant proportion of students are already relatively advanced programmers, but there are also people who are just beginning their adventure with programming and do not even know basic programming issues. Regardless of their knowledge, in many cases the studies end prematurely in failure, which has different grounds [2], [5]. Most universities [1] have adopted a model in which one chosen language is introduced from scratch for all students, assuming that they all start their

[1] Department of Computer Science, Lublin University of Technology, Lublin, Poland; e-mail: adam.kiersztyn.pl@gamil.com, a.kiersztyn@pollub.pl

[2] Department of Mathematical Modelling, The John Paul II Catholic University of Lublin, Lublin, Poland; e-mail: krystyna.kiersztyn@gmail.com

[3] Department of Probability Theory and Statistics, The John Paul II Catholic University of Lublin, Lublin, Poland; e-mail: patrycja.jedrzejewska-rzezak@kul.pl

[4] Department of Computer Science, Lublin University of Technology, Lublin, Poland; e-mail: p.karczmarek@pollub.pl

[5] Department of Electrical & Computer Engineering, University of Alberta, Edmonton, Canada; Department of Electrical and Computer Engineering, King Abdulaziz University, Jeddah, Saudi Arabia; Systems Research Institute PAS, Warsaw, Poland; e-mail: w.pedrycz@ualberta.com

programming education afresh. The same applies to the choice of language (or pseudo-language) for teaching students the basics of algorithmics.

Different approaches are tried in the selection of programming languages used to teach both programming basics and algorithmic basics. In some universities, the same language is used to teach both subjects, often duplicating the content entered. In other centres, two different languages are introduced, in which the relevant content is entered. The choice of languages is often dictated by the preferences of the teaching staff and does not take into account the preferences of those concerned, assuming that they do not have the adequate knowledge to make the right choice.

This paper presents the results of research describing the preferences of students beginning their studies in Computer Science. The starting point for the research was the hypothesis that students starting their studies in computer science know the needs of the labor market and are aware of current trends. Moreover, based on their knowledge, as well as on information obtained from various sources, they are able to deliberately indicate the best languages to teach the basics of programming and the basics of algorythmics. A representative sample (121 for programming basics and 126 for algorithmic basics) of students of the first year of studies in the field of Informatics has been tested. Relevant questionnaires were carried out during the first laboratory classes on the basics of programming and introduction to algorithmics. The research was not limited to a simple survey of student preferences, but used a more complex tool, which is the Analitic Hierarchy Process (AHP) [17], [18].

The work is organised in the following way. Section 2 briefly describes how the AHP method works. Section 3 includes a description of the AHP method modification. Sections 4 and 5 contain analyses of conducted surveys. Section 6 presents discussion of the results obtained. Finally, the last section 7 contains conclusions and description of planned further research.

2. Description of the AHP method

The AHP method [17], [18] is used to solve multi-criteria decision-making problems. The algorithm can be reduced to four essential stages:

- determining the hierarchical structure of the decision problem analysed;
- definition of the decision maker's preferences and determination of importance assessments for particular elements of the hierarchy;
- examining the consistency of preferences;
- creating the final ranking.

An exemplary hierarchy in the decision-making problem, for which the order of five alternatives is established on the basis of 6 criteria, can be found in Figure 2.

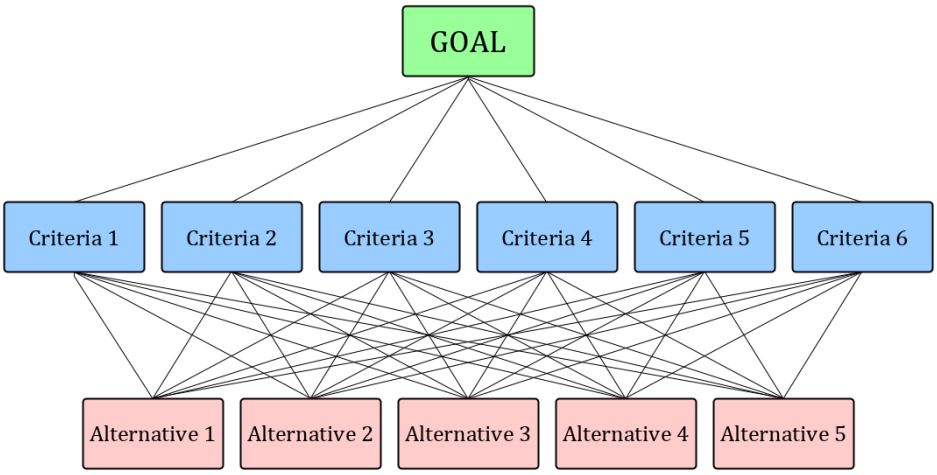


Fig. 2. Example of a hierarchy in a decision-making problem

Source: own study

A special scale is used when determining the decision-maker's preferences (see Table 4) to determine the relationship between the alternatives examined. In each single comparison, two alternatives are analysed in detail and the relationship between them is determined.

Table 4. Linguistic and numerical descriptors used in the AHP method

Descriptor	**Numerical value**
A extremely superior to B	9
A significantly higher than B	7
A exceeds B	5
A slightly higher than B	3
A is equivalent to B	1

Source: own study

The results of the evaluations are recorded in pairs in the form of a square matrix meeting the following consistency criteria:
1. The elements on the diagonal are equal to 1 (this value is completed automatically, as it makes no sense to compare the alternative with itself).
2. The value of the assessment of alternative B relative to A is the inverse of the assessment of alternative A relative to B, i.e. $a_{ij} = 1/a_{ji}$.

For such a matrix, eigenvalues are determined, and then, on the basis of the highest eigenvalue, weights of individual alternatives within a given criterion

are determined. It should be noted here that it is acceptable (although undesirable) that there is some inaccuracy or lack of consistency, in determining the relationship between the individual alternatives. It is possible that alternative A exceeds alternative B, alternative B exceeds alternative C and alternative C exceeds alternative A.

In order to avoid the occurrence of such inconsistencies in determining the relations between the analysed alternatives, author of the method introduced two indicators measuring the level of inconsistency of responses: CI (Consistency Index) and CR (Consistency Ratio). It is assumed that a given matrix is sufficiently consistent if the CR value does not exceed the thresholdof 0.1. However, in some situations values for higher thresholds are acceptable. The CI is determined by the formula

$$CI = \frac{\lambda_{max}-n}{n-1},$$

where λ_{max} means the maximum eigenvalue of the analysed matrix, and n means the number of rows of this matrix. The CR is given by the formula

$$CR = \frac{CI}{RI},$$

where RI stands for the Random Consistency Index and the values can be found in the appropriate tables for different occurrences of n.

In the problems considered, we limit ourselves to a very simple hierarchy in which there is only one criterion with several alternatives. In the case of choosing the best language to learn programming there are 6 alternatives, while in the case of choosing the best language to learn the basics of algorithmics there are 7 alternatives. Moreover, it should be noted that for people who have no experience in indicating dependencies when comparing pairs, it can be very difficult to determine the right numerical value, describing the nature of the relationship. Therefore, due to the lack of experience in determining relationships using the AHP method, it was decided to replace the classic AHP method with its graphic modification.

3. Modification of the classical AHP method

Instead of the classic AHP method, a modification was used in the research, thanks to which the respondents are not required to master complex dependencies of the descriptors used (Table 4). The proposed modification [9] is based on the use of graphical components enabling the selection of values from a certain range. From among many selectable components, a slider was used, which is available on the website [11] used to conduct surveys. The range of values entered is limited to a set of integers $\{0,1,2,\ldots,100\}$ and is represented

by an appropriate component. The transformation of a single answer follows the scheme shown in Figure 3.

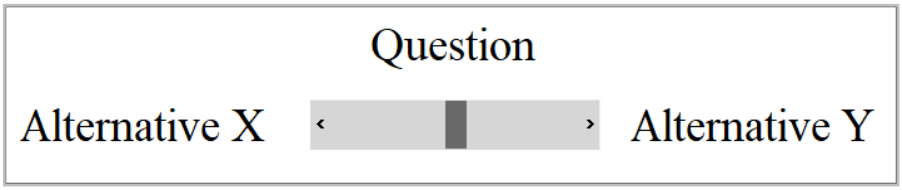


Fig. 3. Slider for comparing the two alternatives X and Y

Source: own study

Answer X is assigned a value of 0, while answer Y is assigned a value of 100. A fixed numerical value of 50 is subtracted from the selected value and then the following conversion is applied

$$f(x) = \begin{cases} 9, & if \quad x < -40 \\ 7, & if \quad x < -30 \\ 5, & if \quad x < -20 \\ 3, & if \quad x < -10 \\ 1, & if \quad x < 10 \\ 1/3, & if \quad x < 20 \\ 1/5, & if \quad x < 30 \\ 1/7, & if \quad x < 40 \\ 1/9, & if \quad x \geq 40. \end{cases} \tag{1}$$

From function (1) we obtain values used in the classical AHP method. For example, if the respondent indicates that he or she prefers the answer "X" by moving the slider handle towards his or her preferred answer and the slider indicates a value of 25, then, according to the above transformations after deducting the constant 50, we obtain the preference "X exceeds Y", which corresponds to the numerical value 5.

4. Analysis of the students' answers

4.1. Analysis of questionnaires concerning the basics of programming

As part of a survey on programming basics, students compared the following languages in pairs, indicating which of them is best suited to learning the programming basics:

- C;
- C++;
- C#;
- Java;
- Pascal;
- Python.

In addition, the students indicated whether and which programming languages they had already learned. Thus, the survey was carried out as shown in Figure 4.

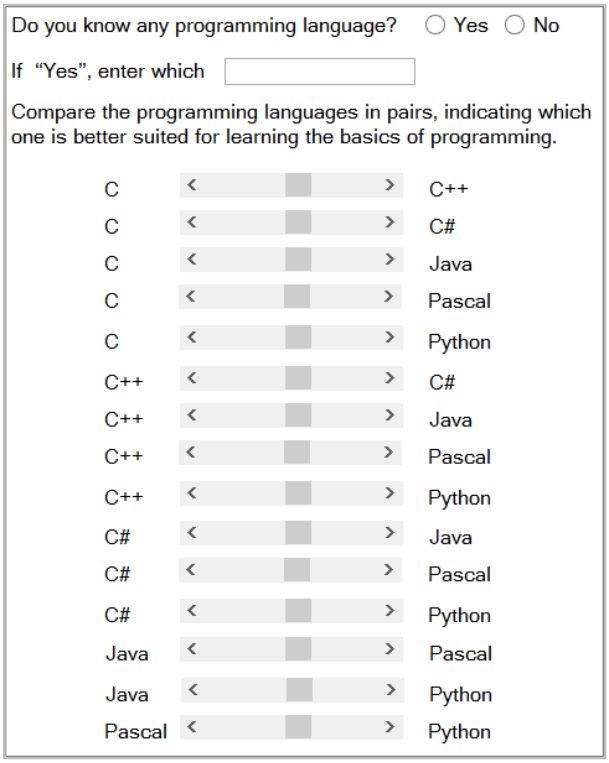


Fig. 4. Survey on choosing the best language to learn the basics of programming

Source: own study

Using the graphical modification of the AHP method, the preferences of individual students were determined and reliability indexes of their responses were defined. Additionally, the reliability index CI^* defined on the basis of CI values was introduced, scaling it linearly to a value from the range $0-1$. The

value one corresponded to the highest reliability of the surveyed student, while the value zero to the lowest reliability of a given student.

The answers of individual students made it possible to determine the preferences of languages best suited for learning the basics of programming. The basic statistics of the weighs values (classification) for individual languages are presented in Table 5, where Q_1 and Q_3 denote quartiles, first and third respectively.

Table 5. Values of basic statistics for weights of particular languages determined by AHP

	C	**C++**	**C#**	**Java**	**Pascal**	**Python**
Min	0.0194	0.0409	0.0194	0.0204	0.0147	0.0194
Max	0.6429	0.6429	0.6117	0.5175	0.4784	0.5716
$\boldsymbol{Q_1}$	0.0704	0.1653	0.0578	0.0714	0.0241	0.0729
Median	0.1359	0.2338	0.0936	0.1252	0.0516	0.1510
$\boldsymbol{Q_3}$	0.1971	0.3458	0.1485	0.2513	0.0936	0.2862
Mean	0.1521	0.2716	0.1253	0.1794	0.0751	0.1965

Source: own study

Analysing the results presented in Table 5 we can conclude that the languages most popular among students are C++, Python and Java. For these languages, the highest number of "wins" among students' judgements was also recorded. The comparison of the values of the best-rated languages with the number of people declaring their knowledge of a given language is shown in Figure 5.

As can be seen, students prefer already known programming languages. In most cases they identified those languages with which they had already been in contact as the best ones to learn the basics of programming. Only in the case of Java is there a significantly reversed relationship: far more people have indicated Java as the best language to learn the basics of programming than were previously in contact with it. Similarly, the number of people who indicated Python as the best language to learn the basics of programming exceeds the number of people who declare knowledge of this language. This fact is probably related to the current popularity of the language among job offers for programmers.

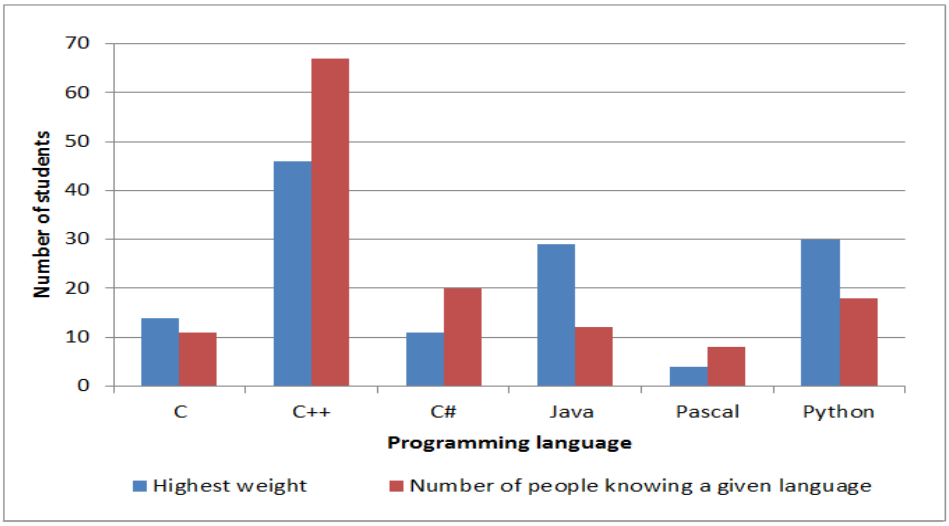


Fig. 5. Comparison of the values of the best-rated languages with the number of people declaring their knowledge of a given language

Source: own study

When we take into account the level of reliability of responses of individual respondents, we get slightly different results. More precisely, the weights assigned to individual languages are different, while the previously established order of weights does not change at all. Only in the case of a total lack of reliability of a given student's answers were all languages assigned a fixed value equal to zero. The values of basic statistics for the weights assigned to individual languages, after applying the reliability index, are presented in Table 6.

Table 6. Values of basic statistics after application of the reliability index CI^*

	C	**C++**	**C#**	**Java**	**Pascal**	**Python**
Min	0	0	0	0	0	0
Max	0.6429	0.6429	0.5769	0.4786	0.3754	0.5454
Q_1	0.0575	0.1169	0.0479	0.0596	0.0187	0.0609
Median	0.0967	0.1983	0.0717	0.0989	0.0380	0.1096
Q_3	0.1667	0.3019	0.1197	0.1668	0.0789	0.2292
Mean	0.1242	0.2253	0.0998	0.1388	0.0617	0.1602

Source: own study

An additional element of the research was to measure the time of filling in the questionnaires by individual respondents. For each student the time of filling in the questionnaires was measured and the obtained result was recorded with an accuracy of one second. With such an additional index, a new method of assessing the reliability of responses of individual respondents was developed. The starting point for building a new reliability index $\omega(t)$ is the assumption that people who solve the test too quickly or too long are not reliable. It was proposed to introduce a transformation based on statistical measures of the position of the sample analysed. A modification of the empirical rule (three-sigma rule) was used, replacing the mean and standard deviation with the median and quartile deviation respectively. The formula realising the given transformation has the form

$$\omega(x) = \begin{cases} 0, & \text{if } t < Me - 3Q \\ 1, & \text{if } t < Me - 2Q \\ 2, & \text{if } t < Me - Q \\ 3, & \text{if } t < Me + Q \\ 2 & \text{if } t < Me + 2Q \\ 1 & \text{if } t < Me + 3Q \\ 0 & \text{if } t \geq Me + 3Q \end{cases},$$

where Me is the median and Q is the quartile deviation.

After multiplying the weights of individual objects obtained by means of the AHP method by the reliability index based on the time of filling in the questionnaire by a given respondent, a new collective classification of objects was obtained. It turned out that by using such a weighting of individual answers, the following classification was obtained: C++ was in the first place among the languages best suited for learning programming according to the group of students analysed, Python was in the second place, while Java, C, C# and Pascal were in the following places. An interesting observation is the fact that such an index is not correlated with the CI-based reliability index. It should be noted, however, that the final classification of individual languages remains the same, regardless of the method of determining the reliability of respondents.

4.2. Analysis of questionnaires concerning the basics of algorithmics

Similar studies have been carried out on students' preferences for choosing the programming language (or pseudo-language) that is best for teaching the basics of algorithmics. A sample of 126 students were examined and asked which of the languages (pseudo-languages) they thought was best for teaching the basics of algorithmics.

The following possibilities were compared in pairs:

- NSD (Nassi–Shneiderman Diagrams);
- C;
- C++;
- C#;
- Java;
- Pascal;
- Python.

The analysis was based on the same considerations as in the case of research on preferences for the best language to teach the basics of programming. The values of basic statistics describing weights assigned to individual languages determined by using the AHP method on the basis of survey results are presented in Table 7.

Table 7. Values of basic statistics for weights of particular languages determined by AHP

	NSD	**C**	**C++**	**C#**	**Java**	**Pascal**	**Python**
Min	0.0067	0.0122	0.0429	0.0073	0.0080	0.0210	0.0352
$\boldsymbol{Q_1}$	0.0975	0.1236	0.1300	0.0467	0.0620	0.0929	0.1071
Median	0.1444	0.1633	0.1914	0.0852	0.0892	0.1527	0.1401
$\boldsymbol{Q_3}$	0.2055	0.2200	0.2333	0.1273	0.1235	0.1972	0.1654
Max	0.4324	0.4271	0.3495	0.3030	0.2769	0.2782	0.3079
Mean	0.1592	0.1770	0.1854	0.0928	0.0997	0.1469	0.1390
Standard dev.	0.0866	0.0826	0.0731	0.0587	0.0525	0.0645	0.0486
Quartile dev.	0.0540	0.0482	0.0516	0.0403	0.0307	0.0521	0.0292

Source: own study

Students' answers, after being processed using the AHP method, are shown in Figure 6. As can be seen, C++ is the most popular language, followed by C, NSD, Pascal, Python, Java and C#.

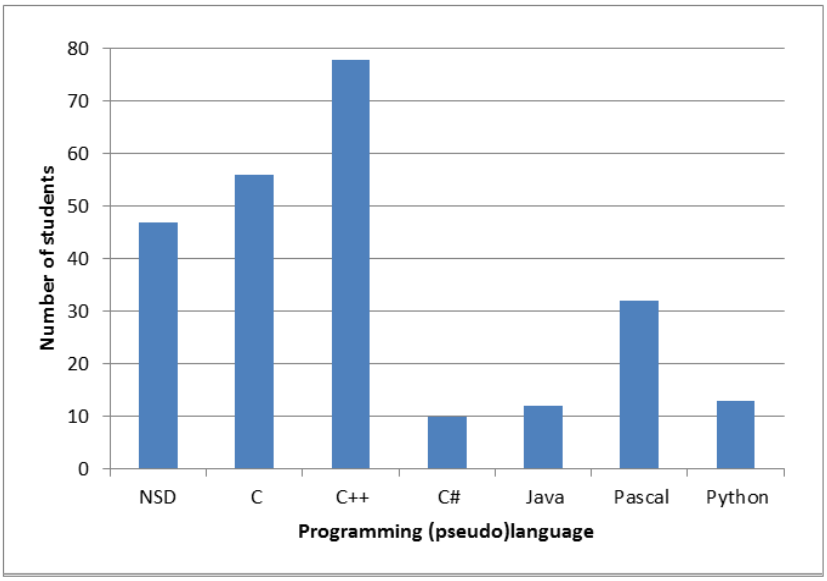


Fig. 6. Number of people indicating a given language as the best to learn the basics of algorithmics

Source: own study

After applying a weighting of the responses using the reliability index described above, the results were similar to the basic results. The basic statistics after taking into account the reliability index are summarised in Table 8.

Table 8. Basic statistics after considering the reliability index CI*

	NSD	C	C++	C#	Java	Pascal	Python
Min	0	0	0	0	0	0	0
Q_1	0.0416	0.0454	0.0576	0.0219	0.0307	0.0485	0.0420
Median	0.0700	0.1000	0.0900	0.0400	0.0500	0.0700	0.0800
Q_3	0.1298	0.1476	0.1327	0.0800	0.0751	0.1126	0.1156
Max	0.3900	0.3800	0.3100	0.1800	0.2000	0.2100	0.2100
Mean	0.0970	0.1074	0.1062	0.0532	0.0559	0.0826	0.0822
Standard dev.	0.0770	0.0748	0.0638	0.0393	0.0376	0.0486	0.0453
Quartile dev.	0.0441	0.0511	0.0375	0.0291	0.0222	0.0320	0.0368

Source: own study

A comparison of the median values for individual languages before and after using the reliability index is shown in Figure 7.

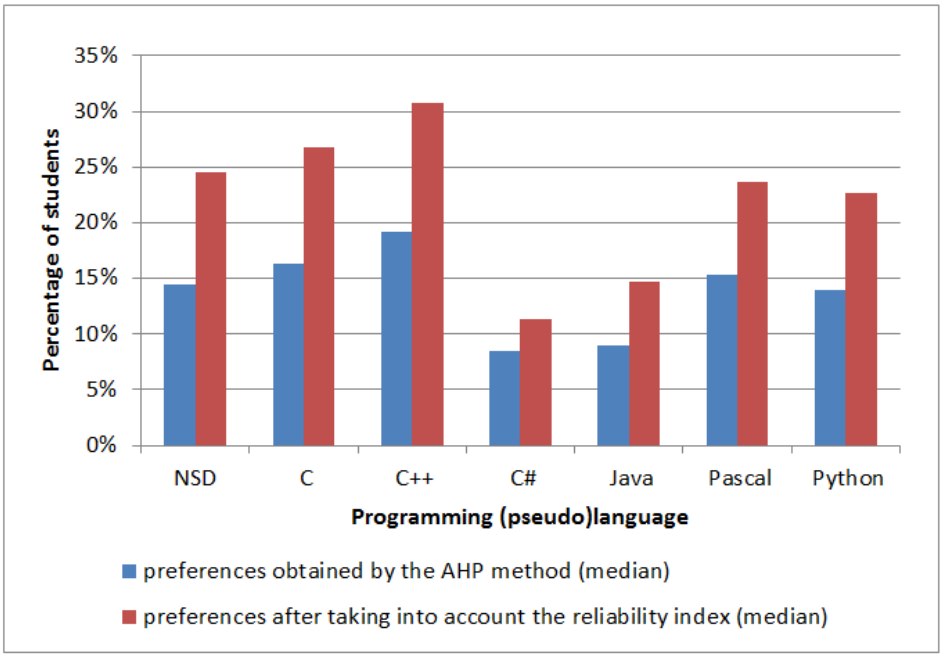


Fig. 7. Comparison of median values for individual languages before and after the application of the reliability index CI^*

Source: own study

5. Discussion

Surveys on both issues analysed were carried out during the first classes among students beginning their studies. Some people solved questionnaires in both subjects, which can be seen in the CR values describing the consistency of respondent's answers. First of all, questionnaires were conducted for the subject of algorithmic basics, then for the basics of programming. Analysing the values of the basic statistical measures for both samples, it can be concluded that the respondents answering questions about the basics of programming showed much more care when filling in the questionnaires and made more thoughtful and consistent answers.

It can be said that each of the measures of position, for a sample that answered questions about the choice of language to learn the basics of programming, is smaller, and thus testifies to more consistent answers. Moreover, the two basic dispersion measures (standard deviation and quartile

deviation) show greater stability of answers, i.e. the whole sample consisted of people with a relatively balanced level of consistency in answering.

The CI values obtained are in most cases not satisfactory (see Table 9), as they significantly exceed a decent level of 0.15. However, due to the impossibility of improving this result, due to the assumption of filling in the questionnaires once, without the possibility of correcting them, and due to the lack of experience of respondents in filling in this type of questionnaires, it was decided not to reject questionnaires with high CI and to use all available data. Thanks to the reliability index, additional weighting of the responses of individual respondents was made, increasing the impact of surveys for which the CI was low and at the same time reducing the impact of the most inconsistent responses on the final result. The use of such a solution made it possible to increase the credibility of the final results. An alternative way to increase the reliability of responses and, consequently, reduce the value of the Consistency Index is to select the thresholds in the formula (1). One solution is to use the Particle Swarm Optimisation (PSO) method [10] to determine optimal threshold values for all respondents.

Table 9. Selected statistics for CI^*

	Basics of algorithmics	**Basics of programming**
Min	0.0595	0
$\boldsymbol{Q_1}$	0.5914	0.1260
Median	1.0039	0.2523
$\boldsymbol{Q_3}$	1.5936	0.4679
Max	2.5643	1.7378
Mean	1.1001	0.3303
Standard deviation	0.6000	0.2930
Quartile deviation	0.5011	0.1710

Source: own study

Moreover, by analysing the time of filling in the questionnaires (with a very similar number of questions), for which the basic statistics are presented in Table 10, we can confirm our conviction that the people answering the questions about the programming basics were more focused on filling in the questionnaires.

Table 10. Values of basic statistics for the time of filling in the questionnaires

	Basics of algorithmics	**Basics of programming**
Min	42	40
Q_1	132	72
Median	168.5	92
Q_3	219.75	123
Max	548	12878
Mean	185.024	203.554
Standard deviation	80.854	1157.491
Quartile deviation	43.875	25.5

Source: own study

It should be noted that, as before, almost all measures of position are smaller for surveys on programming basics. Only for the maximum value we have a deviation from the rule, which resulted from the fact that one person forgot to send the survey immediately after it was completed and did it with a significant delay. As a result, the average (for questionnaires on programming basics), which is very sensitive to outliers, is much higher than the average time of filling in questionnaires on introducing algorithmics.

After removing the outlier, we obtain very stable values of basic statistics (see Table 11).

Table 11. Values of basic statistical measures after removal of an outlier value

Statistical measure	**Basics of programming**
Min	40
Q_1	72
Median	92
Q_3	122.25
Max	185
Mean	97.124
Standard deviation	33.388
Quartile deviation	25.125

Source: own study

Both measures of position and dispersion suggest that the answers given by individual respondents are very thoughtful. Therefore, it should be noted that the results obtained on the basis of questionnaires examining students' preferences for the most suitable programming language for teaching the basics of programming are of greater value.

The obtained results shed new light on the issue of selection of programming languages (pseudo language) for learning the basics of programming and basics of algorithmics by students starting their studies. These results can also be extended to secondary school students, as the research was conducted on students who are just beginning their education at university. Moreover, preferences as to the choice of a programming language were received from the most interested persons, which is undoubtedly a significant added value to the current state of knowledge in this area.

6. Conclusion and future work

The students' preferences for choosing the best language for studying the basics of programming and basics of algorithmics, obtained by means of a well-motivated AHP method, can serve as a suggestion for those who prepare the curriculum for their IT studies. The results of student preferences suggest that C++ is the most popular language, which has an established position and is also desired by many employers. Further positions in the ranking include Java and Python – languages also favoured by employers. This proves a high awareness of the labour market needs of students starting their studies.

It can therefore be concluded that the research confirms the research hypothesis formulated at the beginning. Students starting their studies are aware of the current programming trends and follow the needs reported by employers on an ongoing basis. What's more, they are not afraid to express their views and indicate preferences other than those proposed by the study programs.

It is planned to repeat the research on the same sample in subsequent years in order to be able to determine the trend of changes in students' preferences and to define the impact of broadening knowledge and acquaintance of different languages on their choices. Conducting the research in the following years will allow to verify the research hypothesis that the preferences for choosing a programming language to teach the basics of programming and the basics of algorithmics change with the increase of the respondent's programming awareness.

References

[1] Andrzejewska M., Selected aspects of teaching introductory programming on the computer science studies, Pedagogics, 2016, 25, 75–86.

[2] Andrzejewska M., Factors affecting of computer science students' failure in an introductory-level programming courses, Education-Technology-Computer Science, 2018, 9(4): 211–217.

[3] Beaubouef T., Mason J., Why the high attrition rate for computer science students: Some thoughts and observations, ACM SIGCSE Bulletin, 2005, 37(2): 103–106.

[4] Chatzopoulou D., Economides A., Adaptive assessment of student's knowledge in programming courses, Journal of Computer Assisted Learning, 2010, 26(4): 258–269.

[5] Gomes A., Mendes A., A teacher's view about introductory programming teaching and learning: Difficulties, strategies and motivations, In: 2014 IEEE Frontiers in Education Conference (FIE) Proceedings, 2014, 1–8.

[6] Ivanović M., Budimac Z., Radovanović M., Savić M., Does the choice of the first programming language influence students' grades? In: Proceedings of the 16th International Conference on Computer Systems and Technologies, 2015, 305–312.

[7] Kaczmarczyk L. C., Petrick E. R., East J. P., Herman G. L., Identifying student misconceptions of programming, In: Proceedings of the 41st ACM technical symposium on Computer science education, 2010, 107-111.

[8] Kaplan R. M., Choosing a first programming language, In: Proceedings of the 2010 ACM conference on Information technology education, 2010, 163–164.

[9] Karczmarek P., Kiersztyn A., Pedrycz W., An application of graphic tools and analytic hierarchy process to the description of biometric features, In: International Conference on Artificial Intelligence and Soft Computing, 2018, 137–147.

[10] Kennedy J., Eberhart R., Particle swarm optimization, In: Proceedings of ICNN'95-International Conference on Neural Networks, 1995, 4, 1942–1948.

[11] Koronowicz K., Interankiety, 2011, https://www.interankiety.pl/.

[12] Koulouri T., Lauria S., Macredie R. D., Teaching introductory programming: A quantitative evaluation of different approaches, ACM Transactions on Computing Education (TOCE), 2014, 14(4): 1–28.

[13] Mendes A. J., Paquete L., Cardoso A., Gomes A., Increasing student commitment in introductory programming learning, In: 2012 Frontiers in Education Conference Proceedings, 2012, 1–6.

[14] Parker K. R., Chao J. T., Ottaway T. A., Chang J., A formal language selection process for introductory programming courses, Journal of Information Technology Education: Research, 2006, 5(1): 133–151.

[15] Parker K.R., Ottaway T.A., Chao J.T., Criteria for the selection of a programming language for introductory courses, International Journal of Knowledge and Learning, 2006, 2(1–2): 119–139.
[16] Robins A., Rountree J., Rountree N., Learning and teaching programming: A review and discussion, Computer Science Education, 2003, 13(2): 137–172.
[17] Saaty T.L., What is the analytic hierarchy process? In: G. Mitra, H.J., Greenberg, F.A. Lootsma, M.J. Rijkaert, H.J. Zimmermann (Eds.) Mathematical models for decision support, Springer, Berlin, Heidelberg, 1988, 109–121.
[18] Saaty T. L., Decision making with the analytic hierarchy process, International Journal of Services Sciences, 2008, 1(1): 83–98.
[19] Vihavainen A., Airaksinen J., Watson C., A systematic review of approaches for teaching introductory programming and their influence on success, In: Proceedings of the tenth annual conference on International computing education research, 2014, 19–26.
[20] Vihavainen A., Paksula M., Luukkainen M., Extreme apprenticeship method in teaching programming for beginners, In: Proceedings of the 42nd ACM technical symposium on Computer science education, 2011, 93–98.
[21] Watson C., Li F.W., Failure rates in introductory programming revisited, In: Proceedings of the 2014 conference on Innovation & technology in computer science education, 2014, 39–44.
[22] Xinogalos S., Designing and deploying programming courses: Strategies, tools, difficulties and pedagogy, Education and Information Technologies, 2016, 21(3): 559–588.

Adam Kiersztyn[1], Pavel Urbanovich[2], Nadzeya Shutko[3]

The concept of random cluster based outlier detection

Abstract: Detection of outliers is one of the most common and important problems in modern data analysis. Sources of outliers are different. These could be the result of a database malfunction or user errors. The problem is very important due to the dynamic development of large data sets. Therefore, in this paper we present detailed results of work on the concept of using distribution properties to detect outliers. The aim of the study is to introduce an innovative solution that enables the use of statistical semantics of identification and classification of outliers. The undoubted advantages of the novel approach for outlier detection are the simplicity of interpretation and the possibility of its modification. The effectiveness of the proposed method was compared with other recognized techniques to detecting outliers on both artificially generated and empirical data sets.

Keywords: outlier detection, statistical semantic

1. Introduction

In the realities of the world around us, in every field, especially in biological sciences, we deal with the processing of large amounts of data. Data is growing at an alarming rate, but unfortunately data sets contain outliers. As a result of system malfunction or human error, there are numerous anomalies and outliers in data that can have a dramatic effect on the results of queries, reports, and analyzes performed on such data. Therefore, the process of data integration, in particular the process of detecting and classifying anomalies and outliers, is still under development, and the design of effective anomaly detection methods continues to be mainstreamed in data analysis research.

Researchers have proposed several main directions for working with outliers. They are related to the main areas of work in general machine learning approaches. One of the most important models is based on distance [40], in particular the k-nearest neighbour [7], [8], [26], [37] or the density of a dataset, for example Isolation Forest [31], [32]. Also, tests based on support vector machines [30], [42], hidden Markov models [28], [29] or Gaussian processes [48] have been widely discussed in the literature. Recently, with the increasing popularity of deep neural network applications, such approaches have also emerged carefully analyzed. Models such as self-organizing maps, long-term memory, or convolutional neural networks [12], [13], [33], [50] have been

[1] Department of Computer Science, Lublin University of Technology, Lublin, Poland; e-mail: adam.kiersztyn.pl@gamil.com, a.kiersztyn@pollub.pl

[2] Department of Information System and Technology, Belarusian State Technological University, Minsk, Belarus; e-mail: p.urbanovich@belstu.by

[3] Department of Informatics and Web-design, Belarusian State Technological University, Minsk, Belarus; e-mail: shutko_bstu@mail.ru

widely discussed. Several authors have considered the DBSCAN algorithm, see [43]. Finally, interesting techniques related to fuzzy sets were proposed [11], [16], [18], [21], [22], [36], [47], including fuzzy C-means [19], fuzzy rules [35] or linguistic prototypes [49]. The interested reader can find extensive discussions and method reviews in articles [4], [9], [15], [17]. Recently, there has been extensive research into the use of information granules to detect outliers [5], [10], [14], [17], [20], [24], [25], [51].

The concept described in this paper is based on the use of the distribution properties of the analyzed data. Clusters are randomly generated and the affiliation of individual points to the closest clusters is analyzed. It is reasonable to assume that outliers will not be located near other points.

The article is organized as follows. Section II provides a theoretical description of the proposed innovation. Section III includes detailed numerical experiments on an artificial dataset and two large publicly available databases. Particular attention in this section is devoted to the issue of contextual anomaly detection. Finally, Section IV contains conclusions and further research directions related to the development of the proposed approach.

2. Theoretical description

The starting point of the proposed solution (RCOD) is the use of statistical properties of the analyzed data. In the case of multivariate data, the key element of the analysis is to examine the distribution of the analyzed data.

Suppose we have a set D consisting of N records with K numeric fields each. Such a set can be identified with a matrix with N rows and K columns. For such a data set, we randomly select an *n*-element sample S. We will identify the elements of this sample with the centers of the clusters. The size of the sample should depend on the number of analyzed N elements. In the experimental section, the transformation given by the formula

$$n = \lceil lnN + 1 \rceil,$$

was applied. Where ⌈x⌉is rounding up the value of x. In the next step, the distances between all elements of the sample S are determined. In this way a square table with dimensions n × n is obtained, where the elements on the main diagonal are obviously equal to zero. Then, the distribution of distances between the individual elements of the S sample is analyzed and basic position measures are determined, such as minimum (min S), maximum (max S), quartile1 (Q1 S), quartile 2 (Me S) and quartile 3 (Q3 S) Of course, when determining the minimum value, elements from the main diagonal of the distance matrix are not taken into account. Then, for each element from the input set D, the distances from the centers of the clusters are determined. Information is stored whether the distance to any of the clusters is smaller than the analyzed

statistics. In other words, for each element $x \in D$ from the input data set, vector values are calculated, for which individual components are calculated using the formulas

$$x_{min} = \begin{cases} 1, if\ \ min_{y \in S}\, d(x,y) < min\, S \\ 0, if\ \ min_{y \in S}\, d(x,y) \geq min\, S, \end{cases} \tag{1}$$

$$x_{Q1} = \begin{cases} 1, if\ \ min_{y \in S}\, d(x,y) < Q1\ S \\ 0, if\ \ min_{y \in S}\, d(x,y) \geq Q1\, S, \end{cases} \tag{2}$$

$$x_{Me} = \begin{cases} 1, if\ \ min_{y \in S}\, d(x,y) < Me\, S \\ 0, if\ \ min_{y \in S}\, d(x,y) \geq Me\, S, \end{cases} \tag{3}$$

$$x_{Q3} = \begin{cases} 1, if\ \ min_{y \in S}\, d(x,y) < Q3\, S \\ 0, if\ \ min_{y \in S}\, d(x,y) \geq Q3\, S, \end{cases} \tag{4}$$

$$x_{max} = \begin{cases} 1, if\ \ min_{y \in S}\, d(x,y) < max\, S \\ 0, if\ \ min_{y \in S}\, d(x,y) \geq max\, S. \end{cases} \tag{5}$$

The procedure described above is repeated predetermined number of times M. When subsequent repetitions of the values obtained by the formulas (1–5) are summed. Action proposed solution outlier detection data can be expressed by the following algorithm.

```
For j=1 to M do
  Random n-element sample S.
  Determine the distance between the elements of the set S.
  Determine min(S), Q1(S), Me(S), Q3(S), max(S).
  For i=1 to N do
     Calculate x_min, x_Q1, x_Me, x_Q3, x_max using formulas (1-5).
     Aggregate values x_min, x_Q1, x_Me, x_Q3, x_max
```

The values obtained in this way, describing in how many cases the examined element is located at a certain distance from random cluster centers, allow for the construction of a classifier determining whether a given point can be considered an outlier. Due to the specificity of the analyzed data sets and their significant diversity, it is necessary to develop a dedicated classifier for each set separately.

3. Numerical experiments

The effectiveness of the proposed solution was tested on 4 specially generated two-dimensional data sets and on 26 publicly available empirical data sets: Annthyroid, Arrhythmia, BreastW, Cardio, ForestCover, Glass, Ionosphere, Letter Recognition, Lympho, Mammography, Musk, Optdigits, Pendigits, Pima,

Satellite, Satimage-2, Shuttle, Speech, Thyroid, Vertebral, Vowels, Wbc, Wine [3], [27], [31], [34], [38], [41], [44], [46], [52] coming from the Outlier Detection DataSets (ODDS), and Nad, Unsw0 coming from Kaggle. As part of a series of experiments, the described algorithm was carried out for each data set with the parameter M = 10000. Then, half of the points were randomly selected from the data generated in this way and two classifiers were built on their basis. The first classifier uses Fuzzy Rule (FR) and the second uses Decision Trees (DT). Two well-known measures were used to compare the effectiveness of the proposed method, namely accuracy and precision given by formulas:

$$ACC = \frac{TP+TN}{TP+FP+TN+FN}, \tag{6}$$

$$PREC = \frac{TP}{TP+FP}. \tag{7}$$

The effectiveness of the proposed solution was compared with the classic Isolation Forest method (IF) [31], Gaussian Mixture (GM) [1], Support Vector Machine (SVM) [6], Elliptical Envelope (EE) [40], Local Outlier Factor (LOF) [8]. The characteristics of the analyzed data sets are presented in Table 1.

Table 1. Characteristics of the analyzed data sets

Dataset	**The number of records**	**The number of attributes**
Artificial 1	5090	2
Artificail 2	10400	2
Artificial 3	10600	2
Artificial 4	20400	2
Annthyroid	7200	6
Arrhythmia	452	274
Breastw	683	9
Cardio	1831	21
Cover	286048	10
Glass	214	9
Ionosphere	351	33
Letter	1600	32
Lympho	148	18
Mammography	11183	6
Musk	3062	166
Optdigits	5216	64
Pendigits	6870	16
Pima	768	8
Satellite	6435	36
Satimage-2	5803	36

Dataset	The number of records	The number of attributes
Shuttle	49097	9
Speech	3686	400
Thyroid	3772	6
Vertebral	340	6
Vowels	1456	12
Wbc	378	30
Wine	129	13
Nad	148517	42
Unsw0	257673	43

Source: own study

In the case of generated data sets, an easy-to-interpret graphic visualization of the obtained results is possible. The results for two different aggregation methods are summarized in Figure 1 and Figure 2.

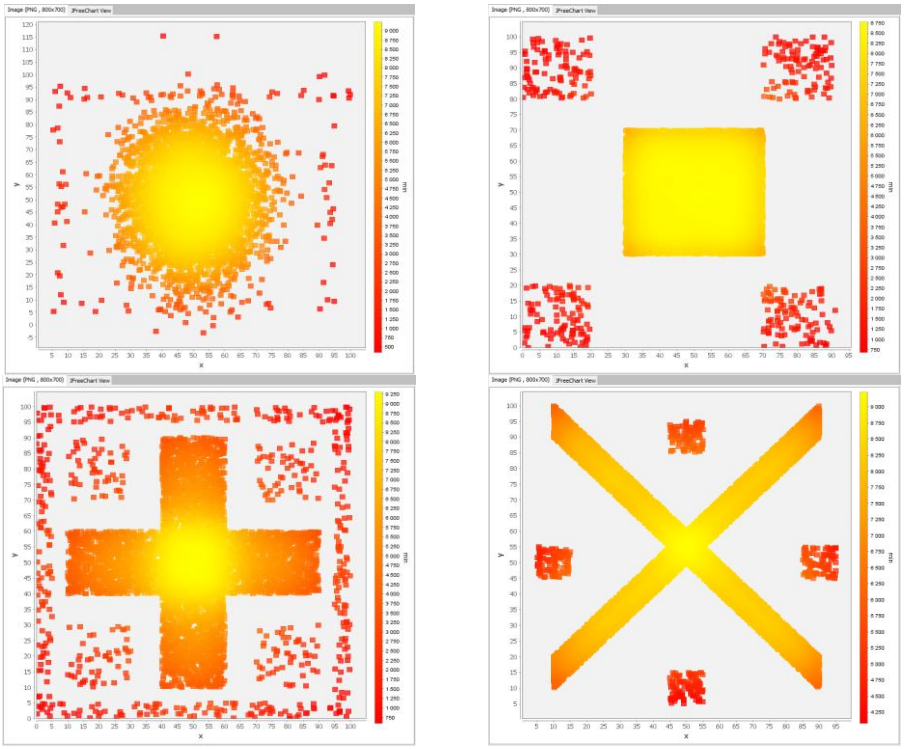


Fig. 1. Values determining the degree of anomaly when applying the minimum function to aggregation

Source: own study

By analyzing the results presented in Figure 1 and Figure 2, it can be stated that the use of the maximum function indicates outliers much more clearly. The differences obtained with this approach are more pronounced. It can be seen that the proper selection of the aggregating function is essential. Moreover, when selecting the aggregating function, one should be guided by the shape of the set and its distribution. When the maximum function is used, fewer elements are designated as outliers. On the other hand, the use of the minimum function determines the elements distant from the center as outliers. It is enough if only one coordinate is sufficiently far from the mean.

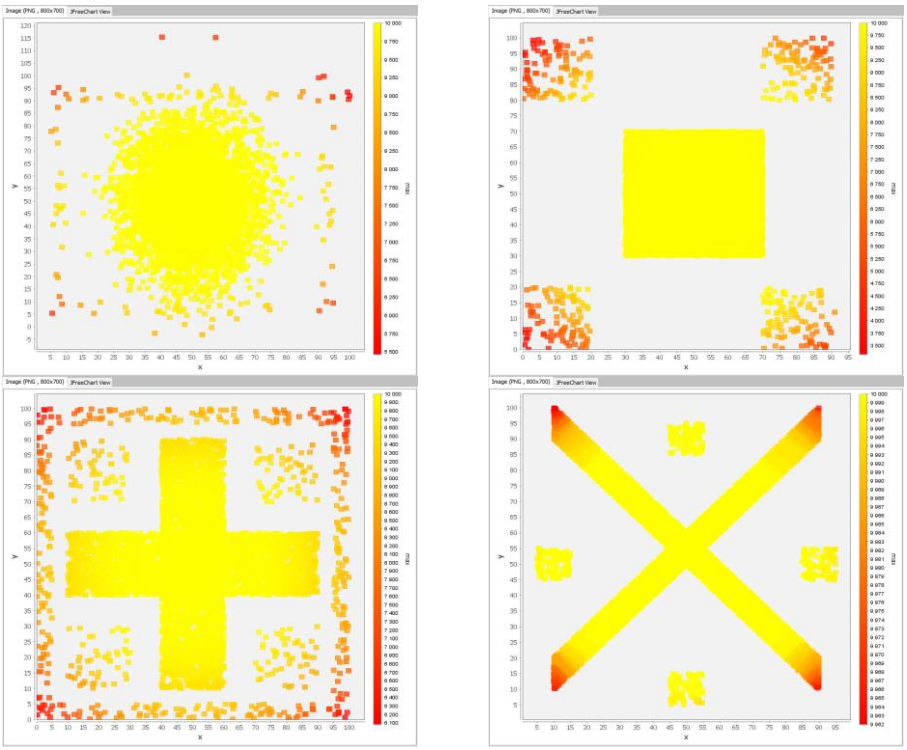


Fig. 2. Values determining the degree of anomaly when applying the maximum function to aggregation

Source: own study

The values of the ACC and PREC measures for the compared methods are presented in the Tables 2 and 3.

Table 2. ACC measure values

Database	**RCOD FR**	**RCOD DT**	**EE**	**GM**	**IF**	**LOF**	**SVM**
Artificial 1	0.993	0.992	0.994	0.993	0.994	0.969	0.992
Artificail 2	1	1	1	1	1	0.929	0.961
Artificial 3	0.999	0.999	0.988	0.987	0.996	0.894	0.924
Artificial 4	1	0.999	0.961	0.961	0.974	0.962	0.977
Annthyroid	0.911	0.885	0.92	0.882	0.899	0.881	0.868
Arrhythmia	0.853	0.845	0.845	0.819	0.843	0.779	0.735
Breastw	0.956	0.9478	0.933	0.933	0.939	0.461	0.388
Cardio	0.979	0.967	0.887	0.809	0.908	0.843	0.891
Cover	0.990	0.988	0.981	0.982	0.982	0.981	x
Glass	0.935	0.954	0.930	0.925	0.930	0.939	0.925
Ionosphere	0.952	0.898	0.917	0.883	0.758	0.863	0.698
Letter	0.932	0.928	0.897	0.911	0.885	0.936	0.884
Lympho	1	0.946	0.959	0.959	0.986	0.973	0.939
Mammography	0.967	0.967	0.954	0.966	0.964	0.958	0.964
Musk	1	1	0.999	0.988	0.998	0.947	0.937
Optdigits	0.992	0.993	0.942	0.943	0.943	0.947	0.946
Pendigits	0.993	0.989	0.959	0.957	0.971	0.958	0.959
Pima	0.611	0.591	0.664	0.655	0.674	0.509	0.609
Satellite	0.893	0.877	0.801	0.656	0.73	0.577	0.673
Satimage-2	0.998	0.998	0.991	0.985	0.996	0.978	0.977
Shuttle	0.934	0.994	0.963	0.98	0.993	0.87	0.913
Speech	0.981	0.975	0.968	0.969	0.968	0.97	0.967
Thyroid	0.969	0.967	0.983	0.964	0.977	0.952	0.96
Vertebral	0.742	0.7833	0.754	0.758	0.758	0.767	0.754
Vowels	0.945	0.967	0.935	0.957	0.944	0.954	0.943
Wbc	0.952	0.9521	0.937	0.939	0.952	0.91	0.913
Wine	0.985	0.969	0.93	0.853	0.86	0.845	0.891
Nad	0.997	0.993	0.542	0.489	0.510	0.533	x
Unsw0	0.676	0.977	0.646	0.645	0.581	0.557	x

Source: own study

Comparing the values of the ACC measure, it can be safely stated that the proposed solution does not differ from the effectiveness of other recognized methods. An in-depth analysis carried out on a large number of databases, consisting of data with different characteristics, allows for a thesis that the proposed method is effective and has the potential for further development.

Table 12. PREC measure values

Database	RCOD FR	RCOD DT	EE	GM	IF	LOF	SVM
Artificial 1	0.765	0.75	0.820	0.809	0.831	0.111	0.767
Artificail 2	1	1	1	1	1	0.078	0.488
Artificial 3	1	1	0.893	0.888	0.965	0.067	0.328
Artificial 4	1	1	0	0	0.332	0.03	0.409
Annthyroid	0.281	0.198	0.459	0.205	0.318	0.199	0.111
Arrhythmia	0.45	0.479	0.459	0.205	0.318	0.199	0.111
Breastw	0.949	0.932	0.904	0.904	0.912	0.23	0.126
Cardio	0.951	0.840	0.411	0.011	0.523	0.182	0.434
Cover	0.282	0.162	0.019	0.053	0.087	0.026	x
Glass	0	0	0.125	0.111	0.125	0.25	0.111
Ionosphere	0.930	0.877	0.888	0.84	0.664	0.81	0.579
Letter	0.267	0.395	0.18	0.283	0.008	0.49	0.071
Lympho	1	0.25	0.5	0.5	0.833	0.667	0.2
Mammography	0.326	0.326	0.008	0.269	0.232	0.093	0.224
Musk	1	1	0.979	0.814	0.969	0.156	0
Optdigits	1	0.911	0	0	0.013	0.087	0.053
Pendigits	0.906	0.761	0.103	0.045	0.353	0.071	0.09
Pima	0.416	0.388	0.519	0.506	0.534	0.296	0.44
Satellite	0.829	0.812	0.685	0.457	0.573	0.332	0.483
Satimage-2	0.973	1	0.629	0.371	0.845	0.114	0.07
Shuttle	0.860	0.961	0.744	0.858	0.949	0.091	0.391
Speech	0	0	0.033	0.05	0.033	0.1	0.016
Thyroid	0.419	0.4	0.656	0.28	0.538	0.022	0.196
Vertebral	0.053	0.167	0	0.033	0.033	0.067	0
Vowels	0.214	0.5	0.06	0.38	0.18	0.327	0.163
Wbc	0.571	0.667	0.429	0.45	0.571	0.19	0.2
Wine	0.833	0.8	0.556	0	0.1	0	0.3
Nad	0.995	0.992	0.542	0.469	0.491	0.514	x
Unsw0	0.950	0.983	0.723	0.722	0.672	0.654	x

Source: own study

The values of the PREC measure indicate, however, that the proposed solution is characterized by high stability in the correct classification of outliers. This is a very important property, especially if you plan to apply fuzzy set-based modifications.

Analyzing the results presented in Tables 2 and 3, it can be concluded that the proposed solution can easily compete with other recognized methods of detecting outliers. A thorough analysis of the considered measures allows to state that in most of the analyzed databases, the proposed solution has the highest values. Only in a few cases the introduced method differs slightly

from other methods. Usually, however, only one compared method is able to achieve measure values better than RCOD. In addition, it should be noted that the proposed algorithm is stable and every time returns the result of the classification, which is not always true in the case of SVM.

4. Conclusion and future work

The proposed solution for detecting outliers uses statistical data semantics and distribution properties. Through the proper selection of parameters classifying the elements as outliers, a tool was obtained, the effectiveness of which is comparable, or even better, than other recognized methods. In the further stages of developing the concept, it is planned to conduct in-depth research on increasing efficiency through more complex classification methods. In addition, it is planned to apply modifications using operations on fuzzy sets, in particular a good effect may be achieved by combining the proposed solution with anomaly detection techniques using information granules [24], [25].

Bibliography

[1] Aitkin M., Wilson G.T., Mixture models, outliers, and the EM algorithm, Technometrics, 1980, 22(3): 325–331.

[2] Abe N., Zadrozny B., Langford J., Outlier detection by active learning, In: Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining, 2006, 504–509.

[3] Aggarwa C.C., Sathe, S., Theoretical foundations and algorithms for outlier ensembles, Acm sigkdd explorations newsletter, 2015, 17(1): 24–47.

[4] Akoglu L., Tong H., Koutra D., Graph based anomaly detection and descryption: a survey, Data mining and knowledge discovery, 2015, 29(3): 626–688.

[5] Albanese A., Pal S.K., Petrosino A., Rough sets, kernel set, and spatiotemporal outlier detection, IEEE Transactions on knowledge and data engineering, 2012, 26(1): 194–207.

[6] Amer M., Goldstein M., Abdennadher S., Enhancing one-class support vector machines for unsupervised anomaly detection, In: Proceedings of the ACM SIGKDD workshop on outlier detection and description, 2013, 8–15.

[7] Angiulli F., Pizzuti C., Fast outlier detection in high dimensional spaces, In European conference on principles of data mining and knowledge discovery, 2002, 15–27.

[8] Breunig M.M., Kriegel H.P., Ng R.T., Sander J., LOF: identifying density-based local outliers, In Proceedings of the 2000 ACM SIGMOD international conference on Management of data, 2000, 93–104.

[9] Chandola V., Banerjee A., Kumar V., Anomaly detection: A survey, ACM computing surveys (CSUR), 2009, 41(3), 1–58.

[10] Chen Y., Miao D., Wang R., Outlier detection based on granular computing, In International conference on rough sets and current trends in computing, 2008, 283–292.
[11] Chimphlee W., Abdullah A.H., Sap M.N.M., Srinoy S., Chimphle, S., Anomaly–based intrusion detection using fuzzy rough clustering, In 2006 International Conference on Hybrid Information Technology, 2006, 329–334.
[12] Chouhan N., Khan A., Network anomaly detection using channel boosted and residual learning based deep convolutional neural network, Applied Soft Computing, 2019, 83, 105612.
[13] De la Hoz E., De La Hoz E., Ortiz A., Ortega J., Martínez-Álvarez A., Feature selection by multi-objective optimisation: Application to network anomaly detection by hierarchical self-organising maps, Knowledge-Based Systems, 2014, 71, 322–338.
[14] Duraj A., Szczepaniak P.S., Ochelska-Mierzejewska J., Detection of outlier information using linguistic summarization, In Flexible query answering systems, 2015, 101–113.
[15] Fanaee H., Gama J., Tensor-based anomaly detection: An interdisciplinary survey, Knowledge-Based Systems, 2016, 98, 130–147.
[16] Gómez J., González F., Dasgupta D., An immune-fuzzy approach to anomaly detection, In The 12th IEEE International Conference on Fuzzy Systems, 2003. FUZZ'03. 2003, 1219–1224.
[17] Habeeb R.A.A., Nasaruddin F., Gani A., Hashem I.A.T., Ahmed E., Imran M., Real–time big data processing for anomaly detection: A survey. International Journal of Information Management, 2019, 45, 289–307.
[18] Hoang X.D., Hu J., Bertok P., A program-based anomaly intrusion detection scheme using multiple detection engines and fuzzy inference, Journal of Network and Computer Applications, 2009, 32(6): 1219–1228.
[19] Izakian H., Pedrycz W., Anomaly detection in time series data using a fuzzy c-means clustering, In 2013 Joint IFSA world congress and NAFIPS annual meeting (IFSA/NAFIPS), 2013, 1513–1518.
[20] Jiang F., Chen Y.M., Outlier detection based on granular computing and rough set theory, Applied intelligence, 2015, 42(2): 303–322.
[21] Karczmarek P., Kiersztyn A., Pedrycz W., Fuzzy set–based isolation forest, In 2020 IEEE International Conference on Fuzzy Systems (FUZZ–IEEE), 2020, 1–6.
[22] Karczmarek P., Kiersztyn A., Pedrycz W., Al E., K–Means–based isolation forest. Knowledge-Based Systems, 2020, 195, 105659.
[23] Keller F., Muller E., Bohm K., HiCS: High contrast subspaces for density-based outlier ranking, In 2012 IEEE 28th international conference on data engineering, 2012, 1037–1048.

[24] Kiersztyn A., Karczmarek P., Kiersztyn K., Pedrycz W., The Concept of Detecting and Classifying Anomalies in Large Data Sets on a Basis of Information Granules, In 2020 IEEE International Conference on Fuzzy Systems (FUZZ–IEEE), 2020, 1–7.
[25] Kiersztyn A., Karczmarek P., Kiersztyn K., Pedrycz W., Detection and Classification of Anomalies in Large Data Sets on the Basis of Information Granules, IEEE Transactions on Fuzzy Systems, 2021.
[26] Knorr E.M., Ng R.T., Tucakov V., Distance-based outliers: algorithms and applications, The VLDB Journal, 2000, 8(3): 237–253.
[27] Lazarevic A., Kumar V., Feature bagging for outlier detection, In Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining, 2005, 157–166.
[28] Li J.G., Hu X.G., Efficient mixed clustering algorithm and its application in anomaly detection. Jisuanji Yingyong, Journal of Computer Applications, 2010, 30(7), 1916–1918.
[29] Li J., Pedrycz W., Jamal I., Multivariate time series anomaly detection: A framework of Hidden Markov Models, Applied Soft Computing, 2017, 60, 229–240.
[30] Lin S.W., Ying K. C., Lee C.Y., Lee Z.J., An intelligent algorithm with feature selection and decision rules applied to anomaly intrusion detection, Applied Soft Computing, 2012, 12(10): 3285–3290.
[31] Liu F.T., Ting K. M., Zhou Z.H., Isolation forest, In 2008 eighth IEEE international conference on data mining, 2008, 413–422.
[32] Liu F.T., Ting K.M., Zhou Z.H., Isolation-based anomaly detection, ACM Transactions on Knowledge Discovery from Data (TKDD), 2012, 6(1): 1–39.
[33] Malhotra P., Vig L., Shroff G., Agarwal P., Long short term memory networks for anomaly detection in time series, In Proceedings, 2015, 89–94.
[34] Micenková B., McWilliams B., Assent I., Learning outlier ensembles: The best of both worlds-supervised and unsupervised, In Proceedings of the ACM SIGKDD 2014 Workshop on Outlier Detection and Description under Data Diversity (ODD2), 2014, 51–54.
[35] Moshtaghi M., Bezdek J. C., Leckie C., Karunasekera S., Palaniswami M., Evolving fuzzy rules for anomaly detection in data streams, IEEE Transactions on Fuzzy Systems, 2014, 23(3): 688–700.
[36] Östermark R., A fuzzy vector valued KNN-algorithm for automatic outlier detection, Applied Soft Computing, 2009, 9(4), 1263–1272.
[37] Ramaswamy S., Rastogi R., Shim K., Efficient algorithms for mining outliers from large data sets, In Proceedings of the 2000 ACM SIGMOD international conference on Management of data, 2000, 427–438.
[38] Rayana S., Akoglu L., Less is more: Building selective anomaly ensembles, ACM transactions on knowledge discovery from data, 2016, 10(4), 1–33.

[39] Rayana S., ODDS Library. Stony Brook University, Department of Computer Sciences, 2016, Available: http://odds.cs.stonybrook.edu.
[40] Rousseeuw P.J., Driessen K.V., A fast algorithm for the minimum covariance determinant estimator, Technometrics, 1999, 41(3): 212–223.
[41] Sathe S., Aggarwal C., LODES: Local density meets spectral outlier detection, In Proceedings of the 2016 SIAM international conference on data mining, 2016, 171–179.
[42] Schölkopf B., Platt J.C., Shawe-Taylor J., Smola A.J., Williamson R.C., Estimating the support of a high-dimensional distribution, Neural computation, 2001, 13(7): 1443–1471.
[43] Scitovski R., Sabo K., DBSCAN-like clustering method for various data densities, Pattern Analysis and Applications, 2020, 23(2): 541–554.
[44] Tan S.C., Ting K.M., Liu T.F., Fast anomaly detection for streaming data, In Twenty-Second International Joint Conference on Artificial Intelligence, 2011.
[45] Ting K.M., Tan S.C., Liu F.T., Mass: A new ranking measure for anomaly detection, Gippsland School of Information Technology, Monash University, 2009.
[46] Ting K.M., Zhou G.T., Liu F.T., Tan J.S.C., Mass estimation and its applications, In Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining, 2010, 989–998.
[47] Tsang C.H., Kwong S., Wang H., Genetic–fuzzy rule mining approach and evaluation of feature selection techniques for anomaly intrusion detection, Pattern Recognition, 2007, 40(9): 2373–2391.
[48] Wang B., Mao Z., Outlier detection based on Gaussian process with application to industrial processes, Applied Soft Computing, 2019, 76, 505–516.
[49] Wilbik A., Keller J.M., Bezdek J.C., Linguistic prototypes for data from eldercare residents, IEEE Transactions on Fuzzy Systems, 2013, 22(1): 110–123.
[50] Zhou C., Paffenroth R.C., Anomaly detection with robust deep autoencoders, In Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining, 2017, 665–674.
[51] Zhu X., Pedrycz W., Li Z., Granular models and granular outliers, IEEE Transactions on Fuzzy Systems, 2018, 26(6): 3835–3846.
[52] Zimek A., Gaudet M., Campello R.J., Sander J., Subsampling for efficient and effective unsupervised outlier detection ensembles, In Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining, 2013, 428–436.

Michał Kuśpit[1], Paweł Karczmarek[2]

A comprehensive experimental comparison of COVID-19 epidemiological models

Abstract: In this study, we thoroughly analyse selected epidemiological models describing the behavior of the spread of COVID-19 in chosen countries. They are Brazil, Czech Republic, Germany, Italy, and Poland. The discussed models are SIR, SIS, and SEIR. The analysis is based on the publically available datasets. The classic models of epidemiology are discussed because of their intuitively appealing mathematical models as well as their efficiency and popularity. In this chapter we recall the theoretical background of these models as well as we present the results of mathematical modelling using associated formulas. Particularly interesting are the difference obtained between the results with respect to countries and used models.

Keywords: epidemiological models, COVID-19, SIR model

1. Introduction

At the turn of 2019 and 2020, people infected with the new the SARS Cov-2 coronavirus were detected for the first time in the city of Wuhan, China, according to some information from bats [1]. At first, the virus did not seem dangerous to people outside of China, because the only officially detected cases of infected people were people in the Hubei province, whose capital is the city of Wuhan. The increase in daily infections throughout China was very quickly observed, which resulted in the introduction of restrictions that limited the transmission of the virus to people. However, the virus made its way outside China and spread around the world. Currently, there are infected people in every country. Some experts claim that inadequate behavior of the rulers of individual countries led to the introduction of the global pandemic announced on March 11, 2020 by the WHO.

Therefore, it is important to use tools that allow us to study the dynamics of the pandemic, carry out as many and systematic screening tests as possible to check whether a person is or has been sick, and collect the data accurately.

The main goal of this study is to compare the methods of using epidemiological models to forecast the dynamics of the pandemic development caused by the COVID-19 disease. The paper presents the epidemiological model of the SIR, SIS and SEIR [4, 7, 11]. Epidemiological models make it possible to determine the duration of an epidemic, the dynamics of development, or the estimation of data on the number of infected, cured, and the number

[1] Michał Kuśpit, Department of Computer Science, Lublin University of Technology, Nadbystrzycka 36B, 20-618 Lublin, Poland; e-mail: michal.kuspit@pollub.edu.pl
[2] Paweł Karczmarek, Department of Computer Science, Lublin University of Technology, Nadbystrzycka 36B, 20-618 Lublin, Poland; e-mail: p.karczmarek@pollub.pl

of deaths. In particular, we are interested in a thorough examination how the models work for the data gathered from the selected countries having various medical infrastructure, level of existence, Gross domestic product, etc. Therefore, we have chosen the following countries to examinate: Brazil, Czech Republic, Germany, Italy, and Poland.

This manuscript is a report of the MSc thesis of the first of its authors.

The rest of the paper is as follows. In Section 2, we discuss various epidemiological models. Section 3 covers numerical experiments while Section 4 is devoted to conclusions and future work.

2. Epidemiological modeling

Here, we discuss the most important properties of the discipline of mathematical theory of epidemics, including the considered in the experimental section epidemiological models.

Let us recall that epidemiology is the study of how disease spreads. It is also the study of health conditions in specific populations [2]. Next, epidemiological modeling is a branch of science involving the use of specially prepared mathematical models supporting the analysis and control of epidemics. The application of epidemiological models for infectious diseases can support the processes of decision-making. In general, relevant state institutions prepare various scenarios of the development of the disease and, according to the results, make decisions that also take into account the analysis of costs incurred by the state in order to introduce certain rules limiting the functioning of the state/country [3].

It is worth to stress that, in this paper, only deterministic models are considered. Deterministic models are models that are described by systems of differential equations. Systems built on the basis of a deterministic model do not have any random element, i.e. the evolution of the entire system depends on the initial parameters (initial conditions).

2.1. SIR model

In this study, we use one of the most basic epidemiological models: The Kermack-McKendrick (SIR) model. It consists of the following three compartments: S – Susceptible, I – Infectious, R – Removed [4], see Figure 1.

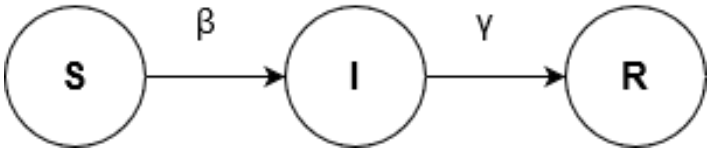


Fig. 1. Diagram of SIR model intervals

Source: [4]

The first interval is the population (S), in the case of this model we assume that S is constant, i.e. as many individuals are born and die. The second range concerns infected people and the size of this group depends on several factors. The first factor is parameter *C*, which describes the average number of contacts of individuals in the population [5], the next factors are the parameters β and γ, they determine how quickly individuals become infected and heal. For the needs of the study, the *C* parameter in the SIR model is constant. The third group is a group of people who contracted the disease and obtained immunity or died (R). The number of units in this group depends on the γ parameter. Thus, the system of differential equations reads

$$\frac{dS}{dt} = -\beta S(t)I(t), \tag{1}$$

$$\frac{dI}{dt} = \beta S(t)I(t) - \gamma I(t), \tag{2}$$

$$\frac{dR}{dt} = \gamma I(t). \tag{3}$$

One estimates the above-mentioned parameters using the L-BFGS-B algorithm with limited memory. The mentioned algorithm is a combination of two subtypes of the BFGS algorithm, i.e. L-BFGS and BFGS-B. L-BFGS-B is an upgrade for the BFGS which is a quasi-Newton method published simultaneously in 1970 by Broyden, Fletcher, Goldfarb and Shanno [6, 10, 12].

The parameters appropriately selected so that the collected data on the number of infected persons in a given area (in our case: A country) from the selected date are adjusted to the SIR model. The fit of parameters (infection) to the model can be verified by counting the residual sum of squares (RSS) [9].

$$RSS = \sum_{i=1}^{n} (\text{Infected} - \text{Fitted})^2. \tag{4}$$

The values of the β and γ parameters range from 0 to 1. In addition, the verification whether the above-mentioned parameters are well selected, i.e.

the data on infection at time *t* are as close as possible to the real data, is checked by counting the standard error (SE)

$$SE = \frac{\sigma}{\sqrt{n}}. \quad (5)$$

2.2. SIS model

The second model which we use in this paper is SIS model. It contains two compartments: S – Susceptible, I – Infectious [7], Figure 2.

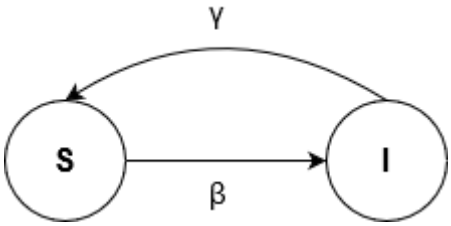


Fig. 2. Diagram of SIS model intervals

Source: [7]

This model is pretty similar to SIR model, but it not contain third compartment (R – Removed). It's caused by that individuals do not get immunity by containment, so the individuals comeback to first compartment which depends from γ factor. The system of differential equations is quite similar to SIR model.

$$\frac{dS}{dt} = -\beta S(t)I(t) + \gamma I(t), \quad (1)$$

$$\frac{\mathrm{dI}}{\mathrm{dt}} = \beta \mathrm{S(t)I(t)} - \gamma \mathrm{I(t)}. \quad (2)$$

The factors β and γ for this model are optimized in the same way as SIR model. Estimation of this parameters were done by using a L-BFGS-B algorithm which is included in optim package in R [9].

2.3. SEIR model

Third model which we use is SEIR model [11]. It's more complicated than those two models, cause it's contains one compartment more. SEIR model contains four compartments: S – Susceptible, E – Exposed, I – Infectious, R – Removed, Figure 3.

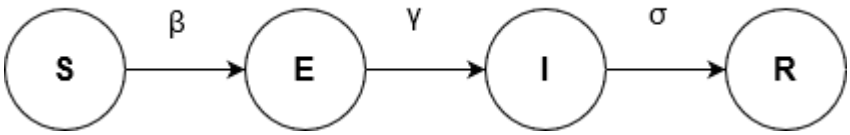


Fig. 3. Diagram of SEIR model intervals

Source: [11]

SEIR model is an upgrade for SIR model. It's contain one more compartment, E – Exposed. Individuals which are sick going to E compartment by a factor β. Exposed compartment shows units which are sick but they don't know about it. They are during the incubation period, they don't have signs that they are sick. Some units from Exposed going to Infected by σ factor which is an incubation rate. Thus, the system of differential equations [8] reads

$$\frac{dS}{dt} = -\beta S(t)I(t), \tag{1}$$

$$\frac{dE}{dt} = \beta S(t)I(t) - \sigma I(t), \tag{2}$$

$$\frac{dI}{dt} = \sigma E(t) - \gamma I(t), \tag{3}$$

$$\frac{dR}{dt} = \gamma I(t). \tag{4}$$

Factor estimation were done in the same way as the other two models. We use L-BFGS-B algorithm which is included in optim package in R [9]. In two models before we optimize two factor β and γ, but in this scenario we need to optimize one more factor σ.

3. Numerical experiments

The research methodology consists in matching the collected data on COVID-19 to the SIR, SIS and SEIR models in several chosen countries: Poland, the Czech Republic, Germany, Italy, and Brazil. For the most countries, we have chosen the dates between 01/06/2021 – 02/24/2021 as the starting point for the model. Based on the data from these days, a model was generated which and then it was fitted to the infection data from the start date to 04/10/2021.

The standard error values for the SIR, SIS, SEIR model for each of the examined countries are presented in Figure 4, Figure 5, Figure 6.

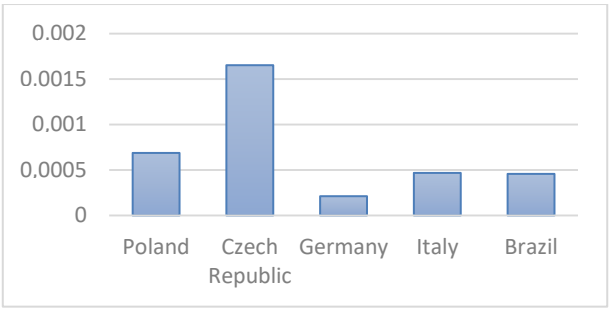


Fig. 4. Standard error values for values from the SIR model

Source: own collaboration

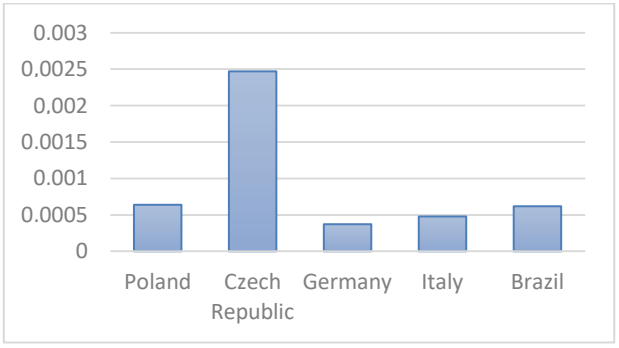


Fig. 5. Standard error values for values from the SIS model

Source: own elaboration

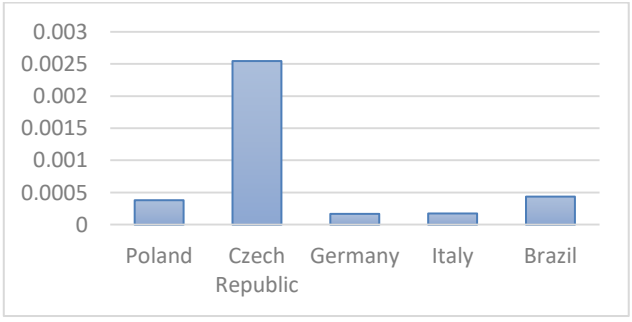


Fig. 6. Standard error values for values from the SEIR model

Source: own elaboration

The starting parameters for each country for the R_0 virus reproduction rate, β virus transmission rate and γ recovery rate are presented in Table 1.

Table 1. Starting parameters for the SIR model

Country	**R_0**	**β**	**γ**
Poland	3.632161	0.014157813	0.003897903
Czech Republic	1.907874	0.033932690	0.017785610
Germany	1.271175	0.032843850	0.025837390
Italy	1.341409	0.053278640	0.039718420
Brazil	1.352930	0.044737010	0.032800560

On the other hand, the values for infections, recoveries and individuals removed from the population (survivors and those who died) on a given model start day are presented in Table 2.

Table 2. Starting values of i(0), r(0), s(0) for the SIR model

Country	**i(0)**	**r(0)**	**s(0)**
Poland	0.04279315	0.03771531	0.91949153
Czech Republic	0.07295602	0.06109213	0.86595185
Germany	0.02892098	0.02754416	0.94353486
Italy	0.04709618	0.04041619	0.91248763
Brazil	0.04839237	0.04440612	0.90720151

The first country we chose to apply the SIR model to the SARS-CoV-2 coronavirus infection data at that time was Poland, see Figure 7.

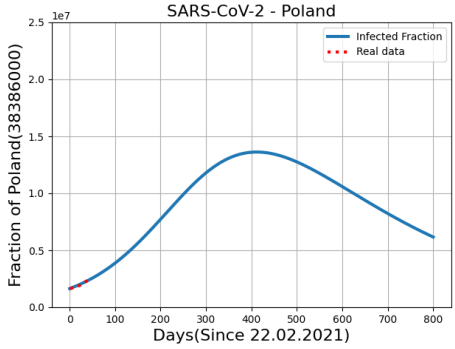


Fig. 7. Curve showing infection data from the SIR model, together with the fit of the collected data in Poland

Source: own elaboration

The next countries that we decided to include in this comparison study are countries such as the Czech Republic, Germany, Italy, and Brazil. It is worth noting that the last one has very high number of infections and the emergence of a new strain of coronavirus – the Brazilian variant (P.1). The results are presented in Figure 8, Figure 9, Figure 10, and Figure 11, respectively.

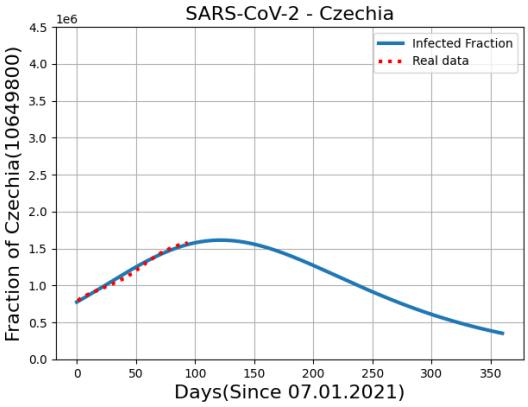


Fig. 8. Curve showing infection data from the SIR model, together with the fit of the collected data in the Czech Republic

Source: own elaboration

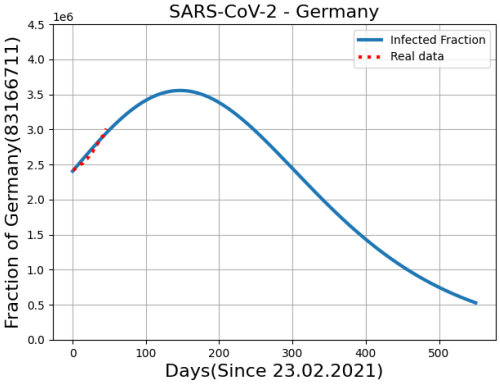


Fig. 9. Curve showing infection data from the SIR model, together with the fit of the collected data in Germany

Source: own elaboration

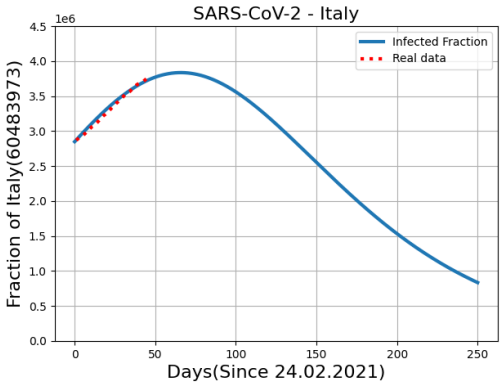


Fig. 10. Curve showing infection data from the SIR model, together with the fit of the collected data in Italy

Source: own elaboration

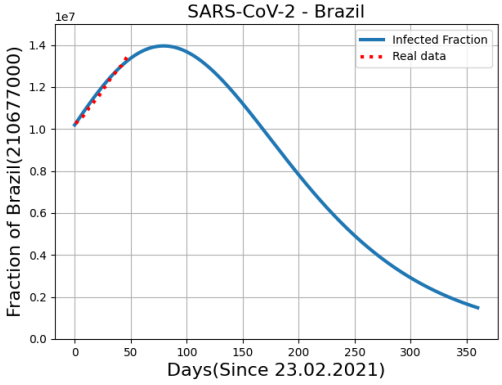


Fig. 11. Curve showing infection data from the SIR model, together with the fit of the collected data in Brazil

Source: own elaboration

By simulating the SIR model for infection values and adjusting the functions to the current data, it was possible to obtain a similar course of the pandemic in a given country. By estimating the maximum number of infections, it is possible to estimate the number of deaths, people requiring hospitalization, or people who need special health care (ventilator therapy, or ECMO therapy), see Figure 12. It is worth noting that one can estimate the percentage of people requiring

hospitalization in a given country at 10% of actively infected people. Consequently, in the case of estimating the number of people requiring intensive care, this figure is 10% of the number of people hospitalized.

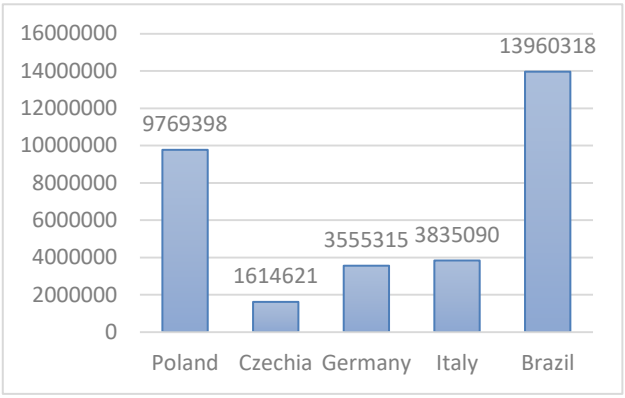


Fig. 12. Maximal number of infections in specific countries

Source: own elaboration

The most recent indicator estimated is the number of deaths. The percentage of people who die is constantly changing and is difficult to quantify. On the other hand, the percentage value is close to around 3% and in the case of this study, we took into account that 3% of people from the entire infected population die. For specific country results, see Figure 13.

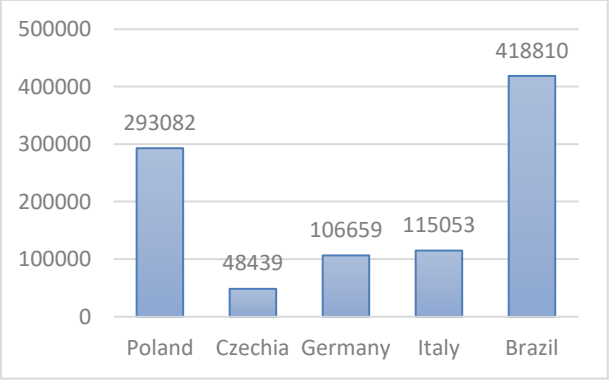


Fig. 13. Maximum number of deaths in individual populations

Source: own elaboration

In order to better illustrate the maximum number of infections in relation to the population of a given country, we calculated the percentage values. Data for individual countries are presented in Figure 14.

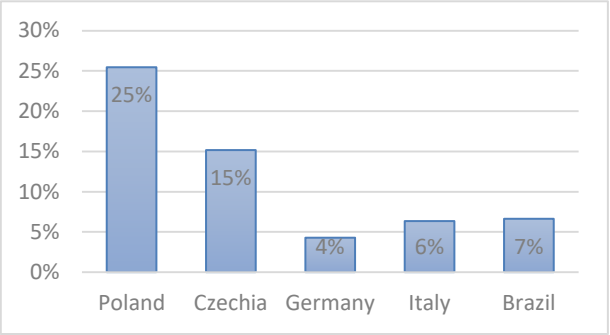


Fig. 14. Maximum number of infections in a particular country (percentage)

Source: own elaboration

The starting parameters for each country for the R_0 virus reproduction rate, β virus transmission rate and γ recovery rate for the SIS model are presented in Table 3.

Table 3. Starting parameters for the SIS model

Country	**R_0**	**β**	**γ**
Poland	1.156511	0.105544850	0.091261430
Czech Republic	1.228799	0.103170700	0.083960600
Germany	1.042245	0.510279400	0.489596500
Italy	1.071282	0.516996400	0.482596100
Brazil	1.072571	0.472656700	0.440676200

Source: own calculations

The starting values for infections, recoveries and individuals removed from the population on a given model are the same as the SIR model. The values are presented in Table 2.

The results of experiment will be presented in the same way as the experiments of the sir model. Firstly we present a SIS model for the Poland, see Figure 15.

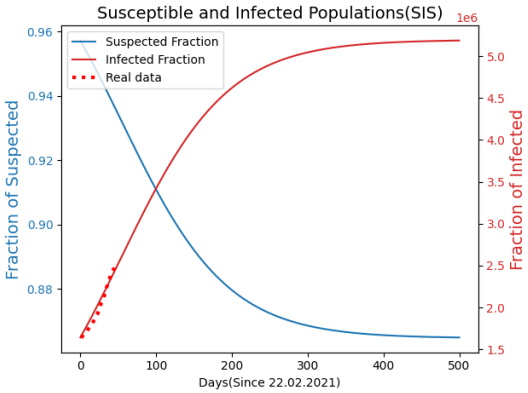


Fig. 15. Chart showing infection data from the SIS model, together with the fit of the collected data in Poland

Source: own elaboration

The rest results of experiment for the next countries will be presented in respectively, Figure 16, Figure 17, Figure 18, Figure 19.

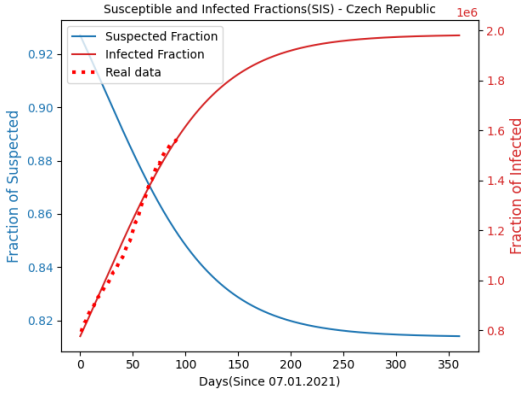


Fig. 16. Chart showing infection data from the SIS model, together with the fit of the collected data in Czech Republic

Source: own elaboration

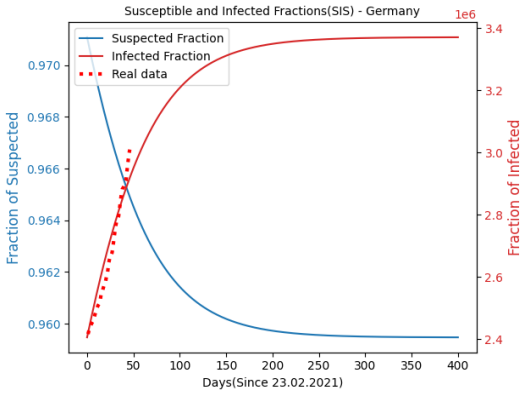


Fig. 17. Chart showing infection data from the SIS model, together with the fit of the collected data in Germany

Source: own elaboration

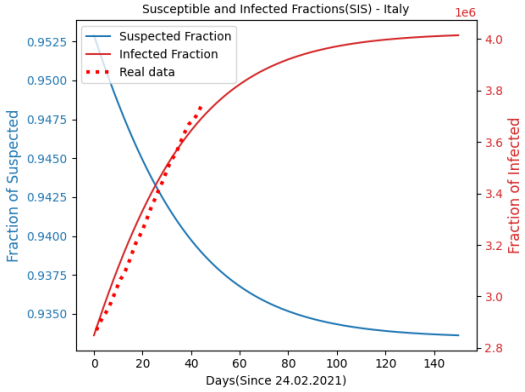


Fig. 18. Chart showing infection data from the SIS model, together with the fit of the collected data in Italy

Source: own elaboration

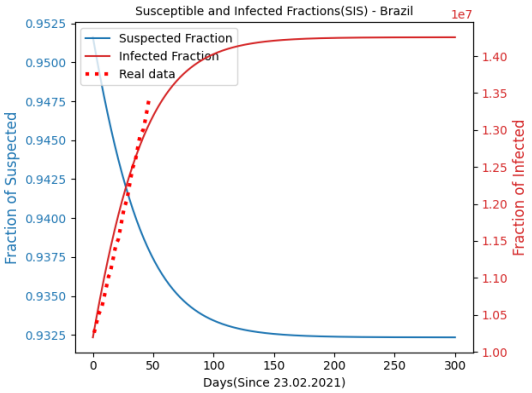


Fig. 19. Chart showing infection data from the SIS model, together with the fit of the collected data in Brazil

Source: own elaboration

By estimating the maximum number of infectious, it is possible to estimate number of deaths and the other values like people who needs a hospitalization. The number of deaths is close to 3% of maximum infectious in each country. The maximum value of infectious estimated by the SIS model is presented in Figure 20, on the other hand maximum value of deaths in particular country is presented in Figure 21.

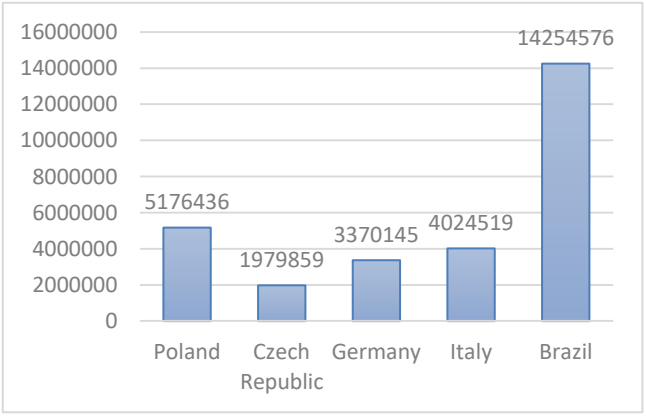


Fig. 20. Maximal number of infections in specific countries (SIS Model)

Source: own elaboration

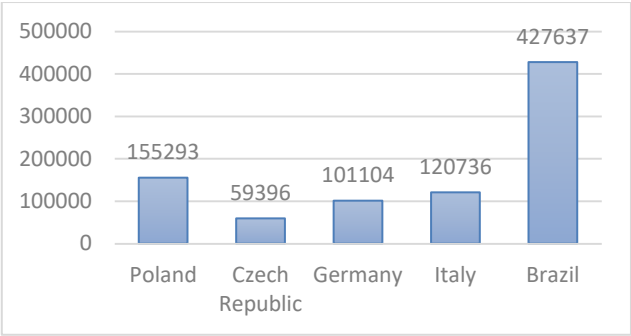


Fig. 21. Maximum number of deaths in individual populations (SIS Model)

Source: own elaboration

For the better illustration of the maximum number of infectious in specific country, we calculated the percentage values, see Figure 22.

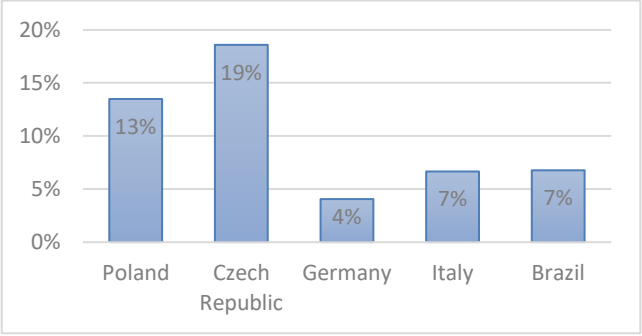


Fig. 22. Maximum number of infections in a particular country (percentage)

Source: own elaboration

Third model which we will present is SEIR model which needs one more parameter compared to SIR and SIS model. It needs a R_0 virus reproduction rate, β virus transmission rate, γ recovery rate and σ incubation rate. The parameters are presented in Table 4.

Table 4. Starting parameters for the SEIR model

Country	**R_0**	**β**	**γ**	**σ**
Poland	4.549779	0.187564700	0.041225020	0.01063322
Czech Republic	6.230057	0.116071828	0.018630943	0.00753181
Germany	1.951506	0.110866230	0.056810610	0.01477168
Italy	2.585964	0.090019020	0.034810620	0.01073313
Brazil	2.291911	0.070535815	0.030775980	0.00987924

Source: own calculations

The starting values for initial conditions in specific country are the same as the other two models presented before. The values are presented in Table 2.

First country which we present for the experiment of the SEIR model will be Poland. Results for the experiment you can see in Figure 23.

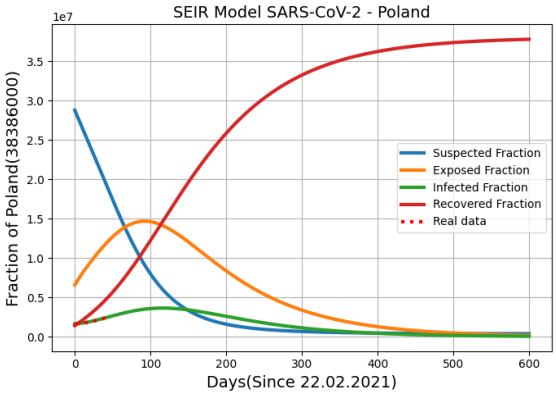


Fig. 23. Chart showing infection, exposed, recovered data from the SEIR model, together with the fit of the collected data in Poland

Source: own elaboration

The next countries that we decided to experiment SEIR model will be presented in the same way as the other models presented before. The order of the presentation for SEIR model depends from the occurrence in Table 4. Next chart for the other countries you can see in respectively in Figure 24 – Figure 27.

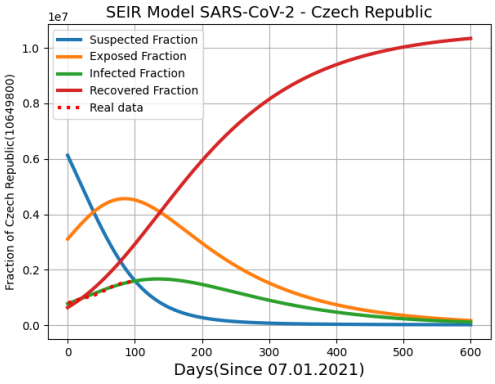


Fig. 24. Chart showing infection, exposed, recovered data from the SEIR model, together with the fit of the collected data in Czech Republic

Source: own elaboration

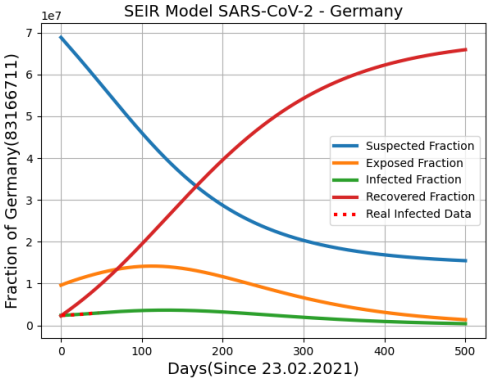


Fig. 25. Chart showing infection, exposed, recovered data from the SEIR model, together with the fit of the collected data in Germany

Source: own elaboration

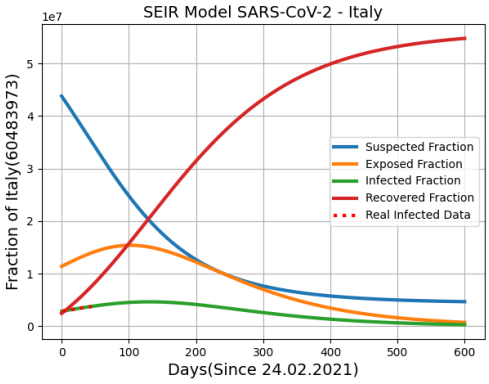


Fig. 26. Chart showing infection, exposed, recovered data from the SEIR model, together with the fit of the collected data in Italy

Source: own elaboration

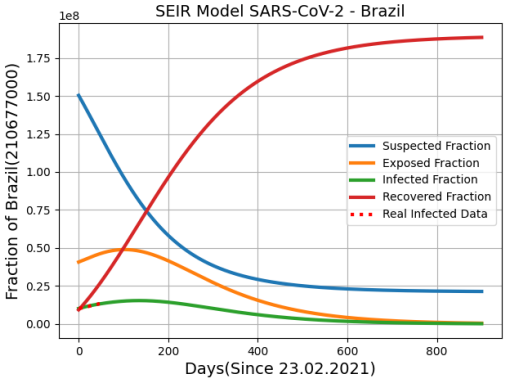


Fig. 27. Chart showing infection, exposed, recovered data from the SEIR model, together with the fit of the collected data in Brazil

Source: own elaboration

The estimation of the maximum number of deaths is estimated in the same way as the other models. The value is a 3% of the maximum value of infectious in particular country, see Figure 29. The number of maximum infectious simulated by the SEIR model you can see at Figure 28.

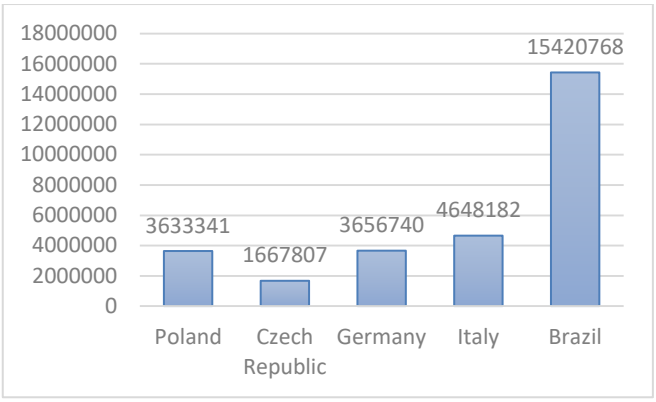


Fig. 28. Maximal number of infections in specific countries (SEIR Model)

Source: own elaboration

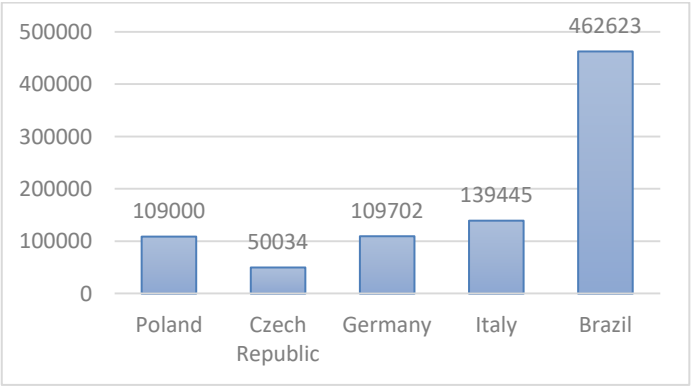


Fig. 29. Maximum number of deaths in individual populations (SEIR Model)

Source: own elaboration

For the better illustration the number of infectious in particular countries, we calculated the percent of infectious individuals in specific countries, see Figure 30.

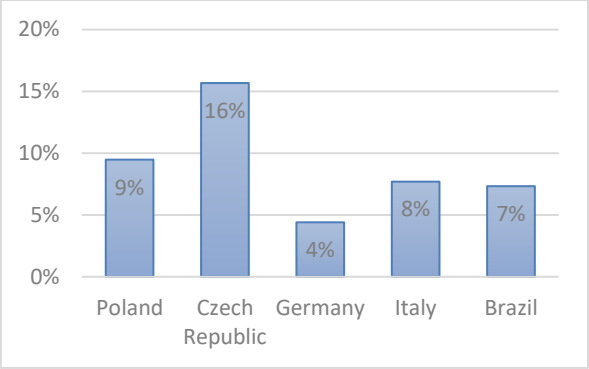


Fig. 30. Maximum number of infections in a particular country (percentage)

Source: own elaboration

4. Conclusions and future work

To sum up, the study includes the results concerning the simulation of the SIR, SIS and SEIR models with the fitting of the current data to the model.

It can be observed from the obtained results that taking into account the maximum number of infectious in relation to the country's population, Germany coped best with the result of 4% of infectious in each of the three models that we tested. On the other hand Poland was the worst country with 25% of infected people from the country's population in case of SIR model. Although SIS and SEIR model worse results of infected people was obtained by Czech Republic with 19% and 16% of infected population. The percentage of cases in Italy is very close to that of Brazil, namely 6% and 7% respectively in case of SIR model. In respect of SIS and SEIR model, percentage of cases in Italy is close to percentage of cases in Brazil as the results in SIR model. The results are 7% to 7% in case of SIS model, and 8% to 7% in case of SEIR model. If we look at the infection curves in Figures 3–7, we can see that the country with the fastest decline in infection rates is Italy in case of SIR model. When it comes to SIS model, in this example, we assume that infected individuals does not obtain immunity, so the countries which coped the best results was Italy and Germany, we can see at Figure 17 and Figure 18. The countries which have a shortest period of wave occurrence were Germany, Brazil and Italy, and the results are pretty close to 150–200 days from the starting point of the model. In case of SEIR model, if we consider a maximum percentage of infectious in specific country the best result was obtained by a Germany, the worse was obtained by Czech Republic with 16% infected population. In respect of SEIR model the country which had the shortest period of epidemic was Czech Republic and is approximately to 300 days from the starting point of model, see Figure 24.

The close result to Czech Republic was obtained by Poland, see Figure 23 with result close to 300–350 days from starting point of SEIR model. However, if we consider deaths, country which has the highest value of deaths was Brazil with value approximately to 462623, see Figure 29. The other countries have pretty close value of deaths in case of SEIR model. Of course, all the obtained results must be approached with a certain degree of uncertainty due to the fact that the all tested models are not fully accurate does not take into account such factor as population variability or vaccination process which goes pretty well in tested counties.

Bibliography

[1] Burki T., The origin of SARS-CoV-2, The Lancet Infectious Diseases, 2020, 20, 1018–1019.

[2] Beaglehole R., Bonita R., Kjellström T., Basic epidemiology, 2nd ed., 2006, World Health Organization.

[3] Jarynowski A., Grabowski A., Modelowanie epidemiologiczne dedykowane Polsce, Portal CZM (Centrum Zastosowań Matematyki), 2015, 1–22.

[4] Kermack W. O., McKendrick A. G., Walker G. T., A contribution to the mathematical theory of epidemics, Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character, 114, 1927, 700–721.

[5] Karp P., Dybiec B., Epidemie a obieg banknotów, Foton 111, 2010, 4–11.

[6] Klawikowska Z., Puchalski B., Skuteczność nowoczesnych algorytmów optymalizacji czerpiących inspirację z procesów naturalnych, Zeszyty Naukowe Wydziału Elektrotechniki i Automatyki Politechniki Gdańskiej 71, 2020, 35–40.

[7] Vargas-De-León C., On the global stability of SIS, SIR and SIRS epidemic models with standard incidence, Chaos, Solitons & Fractals 12, 2011, 1106–1110.

[8] Differential equations for SEIR model. https://docs.idmod.org/projects/emod-hiv/en/latest/model-seir.html. Last accessed: 3.06.2021.

[9] Optimization in R for COVID-19 in Belgium, https://statsandr.com/blog/covid-19-in-belgium. Last accessed: 3.06.2021.

[10] Documentation for optim package for R, https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/optim, Last accessed: 3.06.2021.

[11] Berger D. W., Herkenhoff K.F., Mongey S., An SEIR Infectious Disease Model with Testing and Conditional Quarantine, University of Chicago, Becker Friedman Institute for Economics, Working Paper, 2020, 25.

[12] Comets F., Falconnet M., Loukianov O., Loukianova D., Maximum likelihood estimator consistency for recurrent random walk in a parametric random environment with finite support, Stochastic Processes and their Applications 126, 2014, 3578–3604.

Author Index

Zhanna Alimzhanova	35, 84
Abdyldabek Akkozov	73
Peri Bakasova	9
Katarzyna Baran	95
Szymon Fornal	114
Nella Israilova	9
Adrian Jaczyński	137
Patrycja Jędrzejewska-Rzezak	153
Paweł Karczmarek	114, 137, 153, 182
Adam Kiersztyn	153, 170
Krystyna Kiersztyn	153
Akbota Kulzhanova	48
Michał Kuśpit	182
Marek Miłosz	24
Zainelkhriet Murzabekov	24
Dauren Nazarbayev	35
Gulnara Oruzbaeva	73
Witold Pedrycz	153
Diana Rakhimova	48
Kunduz Sharsheeva	59, 73
Nadzeya Shutko	170
Alima Suleimenova	48
Gulnaz Tultemirova	59, 73
Aliya Turganbayeva	48
Kamshat Tussupova	24
Pavel Urbanovich	170
Salamat Zhunusbayeva	84