AIS 2024

19th International Symposium on

Applied Informatics and Related Areas

PROCEEDINGS

Obuda University

PROCEEDINGS of

19th International Symposium on Applied Informatics and Related Areas

Editor:

Ildikó Petőné Csuka

Obuda University
November 14, 2024
Székesfehérvár, Hungary
ISBN 978-963-449-367-9

CONTENTS

ORAL SESSIONS I-III.

Applied Informatics

Gamifying Cybersecurity: The CTF Challenges

Gergely Vakulya
Alba Regia Technical Faculty
Óbuda University
Székesfehérvár, Hungary
vakulya.gergely@amk.uni-obuda.hu

Helga Anna Albert-Huszár
Alba Regia Technical Faculty
Óbuda University
Székesfehérvár, Hungary
albert.huszar.helga@amk.uni-obuda.hu

***Abstract*—Cybersecurity is becoming increasingly critical in today's digital landscape, yet it remains challenging to enter the field due to its inherently complex nature, requiring expertise across a wide range of specialized areas. This paper explores the use of Capture the Flag (CTF) challenges as a gamified method for enhancing cybersecurity education and skill development. By integrating real-world scenarios into a competitive format, CTF challenges encourage participants to apply their knowledge in dynamic, problem-solving environments. We analyze various CTF challenge categories, and discuss their effectiveness in fostering critical cybersecurity skills.**

***Keywords*—cybersecurity, linux, forensics, cryptography**

I. Introduction

Education is increasingly shifting from traditional classroom settings to virtual environments, necessitating the development of new, interactive modalities to capture learners' attention effectively [1]. As technology advances, educational approaches must evolve to engage students more dynamically, leveraging tools like gamification [2], multimedia contents [3], and collaborative platforms [4]. These methods not only enhance learning experiences but also accommodate diverse learning styles, making education more accessible and effective for a wider audience.

Virtual education environments [5] [6] [7] facilitate real-time interaction between students and educators, promoting active participation and critical thinking. By integrating elements such as simulations, online discussions, and hands-on projects, educators can create immersive learning experiences that foster deeper understanding and retention of material.

One crucial field that warrants focused educational efforts is cybersecurity. As cyber threats become more sophisticated and pervasive, it's essential to equip learners with the skills and knowledge to defend against them. One effective and engaging approach to teaching cybersecurity is through Capture the Flag (CTF) challenges. These competitions provide a hands-on learning experience where participants solve real-world problems and exploit vulnerabilities in controlled environments. [8] [9]

The name of CTF comes fom the traditional game of capture the flag. It involves two teams, each tasked with defending their own flag while attempting to capture the opposing team's flag. In CTF challenges, however, the 'flag' typically takes the form of a hidden or encrypted string (such as a password), which is securely protected. Successfully capturing the flag serves as proof that the player has overcome the challenge by breaching the system's defenses or solving the puzzle.

CTF challenges are based in information technology, often demanding a deep understanding of various technologies that extend well beyond traditional software development. They are designed to simulate real-world security challenges, helping participants understand the tactics used by both attackers and defenders. This practical insight is invaluable in understanding modern security threats.

CTF challenges are typically organized into distinct categories, with the most common being binary exploitation, reverse engineering, cryptography, web technologies, and digital forensics. Each category focuses on a specialized area, requiring participants to apply specific skills and knowledge to solve the associated problems. There are challenges that demand expertise across multiple fields, while others don't fit into any of the afore mentioned categories.

These challenges are part of a CTF event or competition, usually held yearly. While some of them are in-pearson events, the popular ones are online competitions. A CTF competition consist of multiple challenges from varying categories. They are usually team-based, encouraging participants to work together to solve complex problems. Solving challenges can build teamwork and it can be to a team's advantage that each member approaches the same problem differently, not to mention that a team typically consists of members who are good at different challenge categories.

The competitive nature of events keeps participants motivated to improve and learn more as they strive to outperform other teams. Solving complex problems can bring satisfaction, boost confidence and inspire further learning and participation in cybersecurity.

The challenges from past years are available to solve online to develop the skills necessary to participate or to practice for the next competition. There are leaderboards to track the progress as well. As a result it becomes easy to enter the field of cybersecurity.

Unlike traditional classroom learning, CTF challenges provide hands-on experience, making learning more engaging and effective. Some challenges require writing a writeup, which is a document describing in detail the process of solving

the problem. Writeups are important learning resources for people who could not solve that particular challenge or new to cybersecurity. It also broadens one's mind to see how others solved the same challenge.

II. CTF EVENTS

A. Google CTF

Google CTF is an annual, entirely online Capture the Flag (CTF) competition hosted by Google, aimed at both beginner and advanced participants. The event is designed to foster interest in cybersecurity and challenge individuals to solve a variety of technical problems related to hacking, cryptography, and exploitation. It has become one of the most recognized CTF events in the cybersecurity community.

Google CTF features both a beginner-friendly and advanced division, catering to participants with varying levels of expertise. The beginner track, sometimes called "Beginner's Quest," allows newcomers to ease into cybersecurity by solving easier problems. Meanwhile, the advanced track pushes experienced hackers to their limits, tackling complex challenges across multiple disciplines.

B. PicoCTF

picoCTF is one of the largest CTF competitions designed primarily for middle and high school students, but it's open to anyone interested in learning cybersecurity. It was launched in 2013 [10] by Carnegie Mellon University's CyLab Security and Privacy Institute, and continues to be organized annually.

What sets picoCTF apart from other CTF competitions is its educational focus. Unlike many CTFs, which are primarily competitive, picoCTF is designed to teach participants fundamental security concepts while they compete. It features a learning environment called picoGym, where users can practice their skills on challenges from previous years at their own pace. This flexible, learn-as-you-go format is especially useful for beginners and educators looking to integrate cybersecurity into their curricula.

C. Hungarian Cyber Security Challenge

The Hungarian Cyber Security Challenge (HCSC) is Hungary's premier cybersecurity competition, organized annually by the National Cyber Security Center. Being Open to Hungarian citizens aged 16 and above, the event is aimed to identify and promote the next generation of cybersecurity talent.

Unlike most traditional CTF competitions, there are some specific challenges where participants are required to provide detailed documentation, known as "writeups", explaining how they solved the problem and captured the flag. These writeups serve as comprehensive reports that outline the methods and tools used to exploit vulnerabilities and provide step-by-step explanations of the problem-solving process.

Another unique aspect of this event is that the final rankings are not solely determined by the points participants accumulate during the competition. In addition to the scores, the organizers place significant weight on the quality of the participants' writeups and the methods they used to solve the challenges. This approach ensures that creativity, technical depth, and problem-solving strategies are recognized alongside mere point accumulation. By evaluating the techniques and documentation provided in the writeups, the organizers reward participants who demonstrate a thorough understanding of the challenges, innovative approaches, and detailed explanations of their solutions.

D. Riscure Hack Me CTF

As the Internet of Things (IoT) continues to expand, it introduces significant new security challenges, making IoT devices increasingly vulnerable to cyberattacks. IoT ecosystems often include embedded systems with limited security measures, creating potential entry points for malicious actors. To highlight these vulnerabilities and raise awareness of the risks, Capture the Flag (CTF) competitions have started incorporating hardware-based challenges. These challenges focus on the security flaws of embedded systems, which are critical in IoT devices [11]. One standout example is the Riscure Hack Me (rhme) CTF, which ran three times between 2015 and 2018.

III. CHALLENGE CATEGORIES

CTFs are typically organized into specialized challenge types, each focusing on specific areas of cybersecurity. This section outlines the various challenge categories and explains how each one is useful for individuals learning cybersecurity.

A. Web exploration

The web is one of the most ubiquitous components of the cyberworld, forming the backbone of countless online services and interactions [12], [13]. Web-related challenges in CTF competitions often leverage a wide range of technologies, including fundamental elements such as HTML, CSS, and JavaScript, along with various image file formats, frameworks, web servers, and communication protocols. In the simplest challenges, participants might find hidden data embedded in the source code, or they may need to deduce undisclosed file names that contain sensitive information. These types of challenges encourage participants to inspect the structure of web applications and recognize common points of data leakage.

```
GET /search?client=firefox-b-d&q=cat HTTP/3
Host: www.google.com
User-Agent: Mozilla/5.0 (X11; Linux x86_64; rv:130.0) Gecko/20100101 Firefox/130.0
Accept: text/html,application/xhtml+xml,application/xml;q=0.9,image/avif,image/webp,
Accept-Language: en-US,en;q=0.5
Accept-Encoding: gzip, deflate, br, zstd
DNT: 1
Connection: keep-alive
Cookie: ANID=OPT_OUT; SOCS=CAESNQgCEitib3FfaWRlbnRpdHlmcm9udGVuZHVpc2VydmVyXzIwMjQwO
Upgrade-Insecure-Requests: 1
Sec-Fetch-Dest: document
Sec-Fetch-Mode: navigate
Sec-Fetch-Site: none
Sec-Fetch-User: ?1
Priority: u=0, i
TE: trailers
```

Fig. 1. Request header of an example HTTP request.

Other challenges explore the HTTP protocol itself, often requiring participants to manipulate request headers, cookies, or status codes to reveal vulnerabilities or extract data that would otherwise remain concealed. These tasks familiarize

competitors with the nuances of HTTP traffic and the importance of understanding how data is transmitted between client and server. Fig. **??** shows an example HTTP request header.

A third major category involves bypassing authorization mechanisms, where participants exploit flaws in access control or authentication processes. These challenges often simulate real-world vulnerabilities, such as broken authentication, insufficient role checks, or logic flaws that allow unauthorized users to access privileged information or functionalities. By solving these, participants gain critical skills in recognizing and preventing access control vulnerabilities, a common issue in web application security.

B. Cryptography

Cryptography is used in many fields of IT from password or key based authentication to data encryption and validation. Steganography can be considered to be the predecessor of cryptography, but it lives today in a modern form. With software tools arbitrary data (e.g. a string or a data file) can be hidden in an image file. Technically, these programs usually use the least significant bits of the pixels to store the bits. These slight modifications are invisible to the naked eye.

Cryptography plays a vital role in various areas of information technology, ranging from password and key-based authentication to data encryption and integrity validation. It serves as a foundational element for securing sensitive information in our digital world [14].

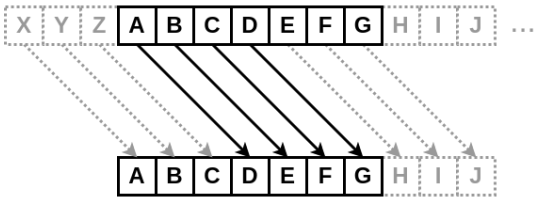


Fig. 2. The method of the Caesar cipher

Cryptography has ancient roots, dating back to antiquity. One of the most basic techniques is the Caesar cipher (see Fig. 2), named after Julius Caesar. This method works by shifting each letter in the plaintext by a fixed number of positions in the alphabet, with the classic Caesar cipher employing a left shift of three. Despite its simplicity, this method laid the foundation for more advanced cryptographic systems. In CTF challenges, basic cryptography problems often begin with similar (e.g. monoalphabetic) ciphers, offering an introduction to fundamental concepts while requiring participants to crack these basic codes through logical analysis and pattern recognition. It is also worth to mention base64, hexadecimal base16 and uu encodings, that are often used to transfer arbitrary binary data through ASCII channels.

More advanced cryptography challenges often exploit vulnerabilities within specific algorithms or their implementations, such as flaws in key generation or usage. Successfully tackling these challenges requires a deep understanding of the mathematical principles underlying the cryptographic techniques. These advanced tasks emphasize not just the mechanics of encryption but also the critical importance of secure implementation in real-world applications.

C. Reverse Engineering

The goal of reverse engineering is to understand how a program or system (such as a network service) operates. One classic task involves figuring out the specific input that will produce a given output. Reverse engineering challenges generally fall into three main categories.

In the first category, participants are provided with the source code, which may or may not be obfuscated. The challenges can be written in any programming language (such as Python, C, JavaScript, or others). In this scenario, participants can modify, compile, and test the code with various inputs, allowing them to experiment with the program's behavior.

In the second category, only the binary of the program is provided. Tools like Ghidra [15] are particularly useful for these challenges, as they can decompile binaries into a pseudo C-like code for analysis (see Fig. 3). After performing static analysis, participants can reorganize the generated code, rename automatically assigned function and variable names, and deduce the program's internal operations. This process helps them understand how the binary works and identify potential ways to manipulate or exploit it.

```
__set_app_type(2);
_DAT_0070f894 = 0xffffffff;
_DAT_0070f898 = 0xffffffff;
puVar1 = (undefined4 *)__p__fmode();
*puVar1 = DAT_0070f88c;
puVar1 = (undefined4 *)__p__commode();
*puVar1 = DAT_0070f888;
_DAT_0070f890 = *(undefined4 *)_adjust_fdiv_exref;
FUN_00409ba1();
if (_DAT_00431410 == 0) {
  __setusermatherr(&LAB_00409b9e);
}
FUN_00409b8c();
_initterm(&DAT_0040b00c,&DAT_0040b010);
local_70 = DAT_0070f884;
__getmainargs(&local_64,&local_74,&local_68,_DoWildCard_0070f880,&local_70);
_initterm(&DAT_0040b000,&DAT_0040b008);
local_78 = *(byte **)_acmdln_exref;
if (*local_78 != 0x22) {
  do {
    if (*local_78 < 0x21) goto LAB_00409b09;
    local_78 = local_78 + 1;
  } while( true );
}
do {
  local_78 = local_78 + 1;
  if (*local_78 == 0) break;
} while (*local_78 != 0x22);
if (*local_78 != 0x22) goto LAB_00409b09;
do {
  local_78 = local_78 + 1;
LAB_00409b09:
```

Fig. 3. Reverse engineering a binary file with Ghidra

The third category involves black-box analysis, where neither the source code nor the binary is provided. Instead, the program is typically accessible through a network connection. Participants must deduce the program's internal workings by systematically providing different inputs and observing the outputs, uncovering vulnerabilities or hidden functionalities purely through external interaction.

D. Binary Exploration

Binary exploitation is a classic technique used to create exploits for binary programs, i.e. to give a specially crafted input, that redirects the flow of the program to call a given function,

```
(gdb) disas main
Dump of assembler code for function main:
   0x0000000000001139 <+0>:     push   %rbp
   0x000000000000113a <+1>:     mov    %rsp,%rbp
   0x000000000000113d <+4>:     sub    $0x30,%rsp
   0x0000000000001141 <+8>:     mov    %edi,-0x24(%rbp)
   0x0000000000001144 <+11>:    mov    %rsi,-0x30(%rbp)
   0x0000000000001148 <+15>:    movl   $0x28,-0x4(%rbp)
   0x000000000000114f <+22>:    movl   $0x75,-0x8(%rbp)
   0x0000000000001156 <+29>:    movl   $0xd6,-0xc(%rbp)
   0x000000000000115d <+36>:    lea    -0x20(%rbp),%rax
   0x0000000000001161 <+40>:    mov    %rax,%rdi
   0x0000000000001164 <+43>:    call   0x1030 <gets@plt>
   0x0000000000001169 <+48>:    nop
   0x000000000000116a <+49>:    leave
   0x000000000000116b <+50>:    ret
End of assembler dump.
(gdb)
```

Fig. 4. Disassembling a binary function in GDB

or execute an attacker code. In these challenges, participants analyze executable files to identify vulnerabilities such as stack or buffer overflows [16] or format string vulnerabilities. Modern compilers and operating systems implement various countermeasures, including stack canaries and address space layout randomization (ASLR), which, in case of particular challenges, must also be bypassed to successfully exploit the program. To aid participants, some challenges also provide the program's source code, offering additional insight into the vulnerability.

Solving these types of challenges typically requires running the program step-by-step, inspecting variables, and examining the state of the stack and registers during execution. One of the most powerful tools for this process is the GNU Debugger (GDB) [17], which allows participants to perform detailed analysis and manipulation of the program during runtime. By using GDB, participants can gain a better understanding of how the binary operates, identify critical points of exploitation, and craft effective payloads to take control of the program. Fig. 4 shows a screenshot of GDB after partially disassembling the main function of the investigated binary executable.

E. Forensics

In real-world digital forensics, investigators analyze seized IT equipment for evidence with uncovering critical information from digital artifacts. Common investigative questions include: Was this particular file or content present on a storage device? Has a file been deleted or altered? What is the true creation or modification date of a file? Who is the genuine sender of a suspicious email? And, was a file transferred from one system to another?

To be able to answer these questions investigators use digital forensics techniques, such as tracking digital footprints, uncovering metadata, and retrieving deleted data. They often examine logs, metadata, file systems, and network traffic. CTF challenges in the Forensics category ask similar questions.

One of the typical challenges in this CTF category involves analyzing a filesystem or a complete storage media image to uncover hidden or deleted data. These images capture the state of the system at a specific point in time. This data is extremely fragile, e.g. booting from the media and interacting with the live system or even mounting the filesystem can inadvertently destroy crucial traces or evidence. Key investigative targets in these cases include uncovering hidden partitions, recovering deleted files or damaged filesystems, and analyzing metadata that might reveal crucial details about the system's activity.

Sometimes, only a few files, such as a webpage or document, need to be analyzed in a forensic challenge. In these cases, the focus is often on examining metadata—hidden information embedded within files that can reveal details like creation dates, modification history, and the tools used to generate the content. Additionally, digital signatures may provide important clues, confirming the authenticity of a file or identifying tampering.

Another possible forensics problem can be connected to steganography. Steganography is the art of hiding data without an actual encryption. This technique allows arbitrary data, such as strings or files, to be hidden within other media, like image files, without raising suspicion. Modern software tools implement steganography by manipulating the least significant bits (LSBs) of pixel values in an image. This method enables data to be embedded in a way that is imperceptible to the naked eye, as the changes made to the pixel values are minimal and often unnoticeable. [18]

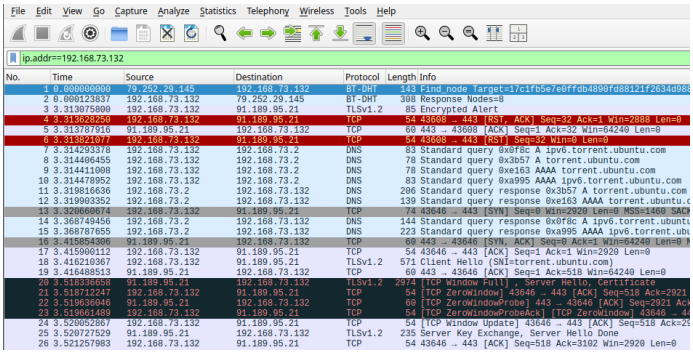


Fig. 5. Analyzing network traffic with Wireshark

A common task in forensics challenges involves analyzing previously captured network traffic. These challenges require participants to dive into packet captures (PCAP files) and interpret network communication data. Understanding network protocols such as TCP, UDP, HTTP, DNS, and others is essential for decoding the traffic.

Tools like Wireshark [19] play a crucial role in this process. Wireshark allows participants to inspect individual packets, filter traffic by protocol or source, and even reconstruct data streams. Fig. 5 shows a screenshot of Wireshark during the analysis of a PCAP capture file. By using this tool, participants can trace the steps of an attacker, find hidden messages, or uncover vulnerable data transmission. Mastering this skillset not only aids in solving CTF challenges but also mirrors the real-world practices of network forensics and incident response, where analyzing network traffic is critical for detecting and responding to cyber threats.

F. Miscellaneous

Miscellaneous challenges in CTFs often encompass problems that don't neatly fit into standard categories. These tasks can be diverse and unpredictable, here will follow some examples, without the need for completeness.

Some challenges involve working with file formats. Participants may be tasked with repairing or modifying corrupted files, such as fixing a broken file header to make the file readable or restorable again. This can include anything from image formats like PNG or JPEG, Solving such problems requires an understanding of the specific file format's structure and how the data is organized within it. Tools like hex editors can be useful in these challenges, as they allow participants to manually inspect and edit the raw bytes of a file.

Many file formats allow for sections that are not directly executed or displayed, creating opportunities to hide additional data or even other files within them. To detect these concealed data, it is essential to recognize key indicators, such as magic bytes—specific sequences of bytes that act as unique identifiers for particular file types.

Binwalk is a powerful tool widely used for automatically analyzing and extracting files from a variety of binary data formats. It excels at detecting embedded file signatures, such as compressed archives or hidden files, and extracting them for further investigation. Its efficiency in identifying these patterns makes it a popular choice in many CTF challenges and forensic investigations.

A special case of such embedding is a polyglot file, which is a single file that can be interpreted as multiple different file types, such as both a PDF and a PNG, each containing distinct content depending on how the file is opened. This dual or multi-interpretation is possible because polyglot files exploit the structure and format of different file types to store more than one set of data within the same file.

A special subject within cybersecurity is IoT (Internet of Things) and embedded security, where unique vulnerabilities arise that are not typically found in traditional computing environments. In addition to the common threats IoT devices are susceptible to specialized attacks such as fault injection and power analysis [20].

Fault injection is typically done by cutting the power supply for very short times, or injecting specially formed clock pulses, making the CPU to skip instructions. With suck a pulse in the right time, security checks can be bypassed [21].

Power analysis involves monitoring the power consumption of a device while it processes cryptographic operations. Attackers can extract sensitive information, such as encryption keys, by analyzing variations in power usage during specific computations.

These scenarios can only be implemented in a low-level environment, and usually a part of a hardware dedicated CTF event.

IV. Summary

Capture the Flag (CTF) challenges offer a highly effective way of developing essential cybersecurity skills by providing hands-on experience in a competitive, gamified environment. This article highlights three notable CTF events. Google CTF, one of the largest and most prestigious, attracts participants from around the globe to solve intricate security problems. PicoCTF, designed for younger participants and beginners, stands out for its accessible introduction to cybersecurity fundamentals. The Hungarian Cyber Security Challenge holds a pivotal role as the premier national competition in Hungary, nurturing local talent and fostering a culture of cybersecurity excellence. Lastly, Riscure Hack Me CTF stands out as a unique competition due to its hardware-focused approach.

The typical CTF challenge categories were thoroughly reviewed, highlighting their distinct educational benefits. Each category targets specific skills and provides practical, hands-on experience that goes beyond theoretical learning. Web exploration challenges expose participants to vulnerabilities within the most ubiquitous platform, the web. Cryptography introduces modern encryption methods, allowing even beginners to engage without prior mathematical knowledge.

Reverse engineering challenges help participants untangle and understand black box systems or obfuscated code by examining their underlying mechanics. Binary exploitation, often considered an advanced topic, involves uncovering classical vulnerabilities such as buffer overflows. In forensics, participants play the role of cyber detectives, investigating cases involving hidden, deleted, or forged data. Lastly, certain challenges that don't fit into the traditional categories were also discussed. These diverse categories offer gateways for learners to apply and refine their cybersecurity knowledge, helping them develop a well-rounded skill set essential for cybersecurity careers.

References

[1] A. Kumar, P. Kumar, and S. C. Basu, "Student perceptions of virtual education: An exploratory study," in *Web-based instructional learning*. IGI Global, 2002, pp. 132–141.

[2] S. Brull and S. Finlayson, "Importance of gamification in increasing learning," *The Journal of Continuing Education in Nursing*, vol. 47, no. 8, pp. 372–375, 2016.

[3] M. J. Nazir, A. H. Rizvi, and R. V. Pujeri, "Skill development in multimedia based learning environment in higher education: An operational model," *International Journal of Information and Communication Technology Research*, vol. 2, no. 11, 2012.

[4] A. Díaz Lantada and C. d. María, "Towards open-source and collaborative project-based learning in engineering education: situation, resources and challenges," *International Journal of Engineering Education*, vol. 35, no. 5, pp. 1279–1289, 2019.

[5] G. Györök and B. Beszédes, "Artificial education process environment for embedded systems," in *9th International Symposium on Applied Informatics and Related Areas - AIS2014*, 2014, pp. 37–42.

[6] A. Sáfár and B. Beszédes, "Educational aspects of a modular power management system," in *AIS 2019 : 14th International Symposium on Applied Informatics and Related Areas organized in the frame of Hungarian Science Festival 2019 by Óbuda University*, 2019, pp. 163–166.

[7] B. Beszédes, "Artificial education process environment for electropneumatic system," in *2023 IEEE 6th International Conference and Workshop Óbuda on Electrical and Power Engineering (CANDO-EPE)*. IEEE, 2023, pp. 175–180.

[8] N. A. A. Rahman, I. H. Sairi, N. A. M. Zizi, and F. Khalid, "The importance of cybersecurity education in school," *International Journal of Information and Education Technology*, vol. 10, no. 5, pp. 378–382, 2020.

[9] F. B. Schneider, "Cybersecurity education in universities," *IEEE Security & Privacy*, vol. 11, no. 4, pp. 3–4, 2013.
[10] K. Zhang, S. Dong, G. Zhu, D. Corporon, T. McMullan, and S. Barrera, "picoctf 2013 - toaster wars: When interactive storytelling game meets the largest computer security competition," pp. 293–299, 2013.
[11] P. Prinetto, G. Roascio, and A. Varriale, "Hardware-based capture-the-flag challenges," in *2020 IEEE East-West Design & Test Symposium (EWDTS)*. IEEE, 2020, pp. 1–8.
[12] S. Calzavara, R. Focardi, M. Squarcina, and M. Tempesta, "Surviving the web: A journey into web session security," *ACM Computing Surveys (CSUR)*, vol. 50, no. 1, pp. 1–34, 2017.
[13] M. Jensen, N. Gruschka, and R. Herkenhöner, "A survey of attacks on web services: Classification and countermeasures," *Computer Science-Research and Development*, vol. 24, pp. 185–197, 2009.
[14] D. Luciano and G. Prichett, "Cryptology: From caesar ciphers to public-key cryptosystems," *The College Mathematics Journal*, vol. 18, no. 1, pp. 2–17, 1987.
[15] C. Eagle and K. Nance, *The Ghidra Book: The Definitive Guide*. no starch press, 2020.
[16] K.-S. Lhee and S. J. Chapin, "Buffer overflow and format string overflow vulnerabilities," *Software: practice and experience*, vol. 33, no. 5, pp. 423–460, 2003.
[17] R. Stallman, R. Pesch, S. Shebs *et al.*, "Debugging with gdb," *Free Software Foundation*, vol. 675, 1988.
[18] T. Morkel, J. H. Eloff, and M. S. Olivier, "An overview of image steganography." in *ISSA*, vol. 1, no. 2, 2005, pp. 1–11.
[19] V. Ndatinya, Z. Xiao, V. R. Manepalli, K. Meng, and Y. Xiao, "Network forensics analysis using wireshark," *International Journal of Security and Networks*, vol. 10, no. 2, pp. 91–106, 2015.
[20] I. Butun, P. Österberg, and H. Song, "Security of the internet of things: Vulnerabilities, attacks, and countermeasures," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 1, pp. 616–644, 2019.
[21] H. Ziade, R. A. Ayoubi, R. Velazco *et al.*, "A survey on fault injection techniques," *Int. Arab J. Inf. Technol.*, vol. 1, no. 2, pp. 171–186, 2004.

Detection of coronavirus disease using feature extraction and XGBoost

Nesibe Yalçın
Engineering Faculty
Erciyes University
Kayseri, Türkiye
nesibeyalcin@erciyes.edu.tr

Gaye Ediboglu Bartos
Alba Regia Technical Faculty
Óbuda University
Szekesfehervar, Hungary
gaye.ediboglu@uni-obuda.hu

Sibel Ünaldı
Engineering Faculty
Bilecik Şeyh Edebali University
Bilecik, Türkiye
sibel.unaldi@bilecik.edu.tr

***Abstract*—Coronavirus is a respiratory disease caused by a virus called SARS-CoV-2. The coronavirus (COVID-19) pandemic has emerged as a serious global threat to all of humanity, creating great pressure on health systems and resulting in quarantine measures in many countries. Early diagnosis of coronavirus is critical for controlling the sources of SARS-CoV-2 infection and preventing its progression. In this study, an XGBoost model is proposed for the early diagnosis of coronavirus. In addition, the effectiveness of the feature extraction process is investigated and evaluated to improve prediction accuracy. Five different performance evaluation metrics have been used. The performance results indicate that the proposed model can be reliably used to detect coronavirus disease.**

Keywords—coronavirus, COVID-19, diagnosis, feature extraction, machine learning, prediction, XGBoost

I. INTRODUCTION

The COrona VIrus Disease 2019 (COVID-19) emerged in China in December 2019, caused by Severe Acute Respiratory Syndrome Corona Virus-2 (SARS-CoV-2). At the end of 2019, the World Health Organization (WHO) announced COVID-19 a global epidemic. As of September 15, 2024, the cumulative total number of COVID-19 cases reported by WHO worldwide is 776,281,230, with 2,275,267 cases in Europe [1]. The cumulative total number of COVID-19 deaths reported by WHO is 7,065,880 as of September 15, 2024 [1]. The symptoms of this disease, including dry cough, shortness of breath, and fever, can range from mild to severe, and in more severe instances, the disease may lead to pneumonia and even death [2]. Diagnosis of the infection can be performed using a test known as Polymerase Chain Reaction (PCR).

The necessity of reliable, rapid, and convenient diagnostic methods encourages the research for alternative approaches in addition to traditional methods based on laboratory tests. Recently, Artificial Intelligence (AI)-driven tools have played a crucial role in addressing various challenges within the health sector, such as epidemic prevention, precision medicine, medical image inspection, and spreading, as well as for prevention and disease detection [3]. AI presents a promising alternative to the estimation of the diagnosis of symptom-based COVID-19. Integration of medical expertise into AI-driven algorithms could serve as a powerful approach to address the problems posed by COVID-19. It also allows for quicker identification of potential cases, especially in resource-limited settings.

Various studies in the literature have explored AI methods for COVID-19 prediction. In the study [4], an affordable and simple method was presented for the detection of COVID-19. The used dataset contains four main categories of features that can be easily accessed at home. AdaBoost, XGBoost, and the other AI methods were used for the prediction of COVID-19. Accuracy, specificity, sensitivity, mean squared error, Kappa index, Matthews Correlation Coefficient (MCC), and Jaccard score were used to show the performance of the classification. The best classification performance in terms of accuracy (73.31%) was achieved by AdaBoost. In [5], Machine Learning (ML) classification models have been developed to predict the presence of COVID-19 based on 21 symptom-based features. An open-source dataset has been utilized for this purpose. The decision tree algorithm demonstrated the highest effectiveness, with an accuracy of 98.57%, a perfect sensitivity of 1.0, and a strong specificity of 0.97. Another study [6] introduced a novel data mining system that integrates ensemble feature selection techniques with ML classifiers to effectively identify COVID-19 infections. In the presented technique (EFS-MLC system); chi-square, Recursive Feature Elimination, Particle Swarm Optimization, Genetic Algorithm, and Random Forest (RF) based ensemble feature selection methods were used and the best features were extracted considering both COVID-19 datasets used. Ref. [7] presented a deep learning technique that designed XGBoost-SIRVD-LSTM model to estimate COVID-19 infection cases. The model integrates XGBoost for feature selection and combines the SIRVD epidemic model with Long Short-Term Memory (LSTM) for improved estimation. The presented model is tested on two COVID-19 datasets and R^2 is obtained as 0.99 for a single day. In our previous work [8], we conducted a study to analyse and predict the presence of COVID-19 based on symptoms and features. In order to achieve this, we utilized Gated Recurrent Unit (GRU) and LSTM deep learning algorithms, as well as K Nearest Neighbour, Support Vector Machine, Logistic Regression, and MultiLayer Perceptron Neural Network ML algorithms. The best performance was achieved by the GRU, with an accuracy of 98.65%.

This paper introduces an XGBoost model taht is built to estimate the diagnosis of COVID-19 cases. Five different performance evaluation metrics have been used. The effectiveness of the feature extraction is explored, Principal Component Analysis (PCA) and SelectKBest techniques are applied. PCA has presented a higher predictive accuracy of 98.16% and an F1-score of 98.88% compared to SelectKBest with fewer features.

The remaining sections of the study are structured as follows: Section II introduces the used dataset in the study. In Section III, the suggested methodology is detailed. Section IV presents the experimental results and analyses models comparatively. Section V concludes the paper.

II. DATASET

Our models were tested on a dataset titled “Symptoms and COVID Presence,” which is publicly available on Kaggle [9]. It contains information from April 17, 2020, to August 29, 2020, and was updated on August 29, 2020. The used dataset consists of 20 features that represent various symptoms, along with 1 class feature indicating whether a person has COVID-19. Data distribution is given in Table I. The presence of

COVID-19 in potential patients is categorized as "no" or "yes."

TABLE I. DATA DISTRIBUTION

Category	*Count*	*Percentage*
Total Examples	5,434	100%
COVID-19 Positive Patients	4,383	80.7%
Healthy Individuals	1,051	19.3%

COVID-19 affects individuals in different ways, with patients experiencing a range of symptoms. In addition to common symptoms (dry cough, shortness of breath, and fever, etc.), some infected individuals report muscle aches, fatigue, and loss of taste or smell (anosmia) [10].

III. METHODOLOGY

This work focuses mainly on predicting the symptom and feature based diagnosis of COVID-19. The binary classification process using XGBoost models with different feature extraction methods is illustrated in Fig. 1. The process begins with the input COVID-19 dataset. It includes key steps for feature extraction, model training, hyperparameter tuning, model performance evaluation, and classification.

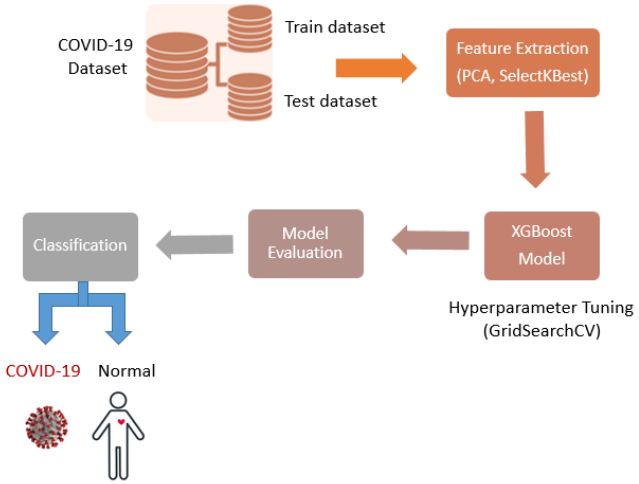


Fig. 1. XGBoost-based classification process for COVID-19 detection

A. PCA

Feature extraction is a useful way to improve model performance and avoid/mitigate overfitting. PCA, first formulated by Karl Pearson [11], [12], is a statistical technique commonly used for data reduction, feature extraction, and classification in machine learning. It generates a new dataset from the original features, consisting of principal components. This technique captures the correlation structure among the features. The first principal component is a linear combination of the features with maximum variance in the dataset, so it contains the most information. The second principal component captures the next maximum variance [13].

B. SelectKBest

SelectKBest technique [14] is a statistical method widely used to extract/select the most relevant features from features in a dataset. This technique applies various scoring functions such as ANOVA F-value, χ^2 statistics, and others, and determines the top K features with the highest scores.

C. XGBoost

XGBoost [15] is a scalable and highly effective machine learning method. This method, based on Gradient Boosting algorithm, applies additional techniques like automatic feature selection and randomization [16], [17]. These techniques help XGBoost resist overfitting, making the model more efficient and accurate.

D. Performance Evaluation

Confusion Matrix (CM), accuracy, recall, precision, F1, MCC, and determination coefficient R^2 metrics were utilized to evaluate the model performance.

IV. EXPERIMENTAL RESULTS

The XGBoost models were developed in Python and implemented using the resources provided by Google Colaboratory. In the experiments, the dataset was first divided into two subsets: 80% for training and 20% for testing. PCA and SelectKBest techniques were used for feature extraction, and the different number of features/components were tested. In the study, the calculation of ANOVA F-value was preferred as a scoring function in SelectKBest technique.

GridSearchCV tool in scikit-learn was used to perform hyperparameter fine-tuning. All combinations of the XGBoost parameters' values presented in Fig. 2 were considered to achieve the best model performance. "learning_rate" determines the step size used to update the model weights at each iteration. A number of estimators (n_estimators) indicate the total number of decision trees created during model training. "max_depth" parameter refers to the maximum depth of each tree. "colsample_bytee" controls the percentage of features randomly sampled for each tree.

```
'learning_rate':(0.01, 0.1)
'n_estimators': (300, 350, 400)
'max_depth': (3, 4, 5)
'colsample_bytree' :(0.2, 0.3)
```

Fig. 2. XGBoost parameters

In addition to GridSearchCV (using 5-fold cross validation) results, the performance of the XGBoost models using PCA and SelectKBest for feature extraction was summarized in Table II and Table III, respectively. As shown in Table II, the best results when PCA was applied were an accuracy of 98.16%, precision of 99.10%, and F1 of 98.88%, achieved with 5 components. On the other hand, no significant change in accuracy or other metrics was observed for 5 or more components. These results show that the first few principal components (especially, up to 5) already captured most of the important information in the dataset.

As seen in Table III, when SelectKBest was applied, the performance of the XGBoost model improved with the increase in the K value. The best performance results were obtained as 98.16% accuracy, 99.1% precision, and 98.88% F1 for 20 features. The results for 5 features were obtained as 94.85% accuracy, 96.96% F1, and 94.2% precision, and PCA+XGBoost achieved higher performance.

TABLE II. PCA + XGBOOST PERFORMANCE RESULTS

Number of Components	XGBoost Best Parameters	Evaluation Metrics						
		CM	*Accuracy*	*Precision*	*Recall*	*F1*	*MCC*	R^2
2	learning_rate: 0.1 max_depth: 3 colsample_bytree: 0.2 n_estimators: 350	$\begin{bmatrix}179 & 13\\ 14 & 881\end{bmatrix}$	0.9752	0.9855	0.9844	0.9849	0.9148	0.8292
5	learning_rate: 0.1 max_depth: 3 colsample_bytree: 0.2 n_estimators: 350	$\begin{bmatrix}184 & 8\\ 12 & 883\end{bmatrix}$	0.9816	0.9910	0.9866	0.9888	0.9373	0.8735
7, 10, 13, and 15	learning_rate: 0.1 max_dept: 3 colsample_bytree: 0.2 n_estimators: 300	$\begin{bmatrix}184 & 8\\ 12 & 883\end{bmatrix}$	0.9816	0.9910	0.9866	0.9888	0.9373	0.8735

TABLE III. SELECTKBEST + XGBOOST PERFORMANCE RESULTS

K	XGBoost Best Parameters	Evaluation Metrics						
		CM	*Accuracy*	*Precision*	*Recall*	*F1*	*MCC*	R^2
3	learning_rate: 0.1 max_depth: 3 colsample_bytree: 0.2 n_estimators: 300	$\begin{bmatrix}137 & 55\\ 50 & 845\end{bmatrix}$	0.9034	0.9389	0.9441	0.9415	0.6646	0.3358
5	learning_rate: 0.1 max_depth: 3 colsample_bytree: 0.2 n_estimators: 300	$\begin{bmatrix}137 & 55\\ 1 & 894\end{bmatrix}$	0.9485	0.9420	0.9989	0.9696	0.8161	0.6478
7	learning_rate: 0.1 max_dept: 3 colsample_bytree: 0.2 n_estimators: 300	$\begin{bmatrix}158 & 34\\ 4 & 891\end{bmatrix}$	0.9650	0.9632	0.9955	0.9791	0.8764	0.7596
10	learning_rate: 0.1 max_dept: 3 colsample_bytree: 0.3 n_estimators: 400	$\begin{bmatrix}169 & 23\\ 13 & 882\end{bmatrix}$	0.9669	0.9746	0.9855	0.9800	0.8842	0.7723
13	learning_rate: 0.1 max_dept: 3 colsample_bytree: 0.2 n_estimators: 400	$\begin{bmatrix}184 & 8\\ 13 & 882\end{bmatrix}$	0.9807	0.9910	0.9855	0.9882	0.9344	0.8672
20	learning_rate: 0.1 max_dept: 3 colsample_bytree: 0.2 n_estimators: 400	$\begin{bmatrix}184 & 8\\ 13 & 882\end{bmatrix}$	0.9816	0.991	0.9866	0.9888	0.9373	0.8735

V. CONCLUSION

This study aims to develop a more efficient and less prone to overfitting XGBoost model by focusing on the important features. PCA and SelectKBest were employed to demonstrate the effectiveness of feature extraction in improving model performance. PCA achieved optimal performance with 5 components. SelectKBest further improved model performance as the number of selected features increased. PCA presented higher predictive accuracy compared to SelectKBest with fewer features. In addition, the obtained results revealed the effectiveness of XGBoost in diagnosing COVID-19. Therefore, combining AI with the other methods will enable faster and more accurate diagnoses. Future research could explore additional feature extraction techniques and their combinations to further improve XGBoost model performance.

REFERENCES

[1] WHO (2024). WHO Coronavirus (COVID-19) Dashboard 2024. Available online at: https://data.who.int/dashboards/covid19/cases (Accessed on Oct. 1, 2024)

[2] B. Yildiz, E. Haspolat, A. Yilmaz, and Z. Önes, "Stability analysis and numerical simulation of dynamic and fractional SEIRD models for spread of nCOVID-19 in Turkey," Asian-European Journal of Mathematics, vol. 15, no. 12, pp. 2250226, 2022.

[3] C. Carmela and C. Pizzuti, "Artificial intelligence for forecasting and diagnosing COVID-19 pandemic: A focused review," Artificial intelligence in medicine vol. 128, (2022): 102286.

[4] A. R. Shabrandi, A. Rajabzadeh Ghatari, N. Tavakoli, M. Dehghan Nayeri, and S. Mirzaei, "Fast COVID-19 Infection Prediction with In-House Data Using Machine Learning Classification Algorithms: A Case Study of Iran. Journal of AI and Data Mining," vol. 11 no. 4, pp. 573-585, 2023.

[5] A. Samad and M. K. Uçar, "Early and Rapid COVID-19 Diagnosis Using a Symptom-Based Machine Learning Model," International Journal of Innovative Science and Research Technology (IJISRT), vol. 9, no. 7, pp. 1537-1543, 2024. Doi: 10.38124/ijisrt/IJISRT24JUL1493

[6] M. J. Hossen, T. T. Ramanathan, A. Al Mamun, "An Ensemble Feature Selection Approach-Based Machine Learning Classifiers for Prediction of COVID-19 Disease," Int. J. Telemed. Appl., vol. 2024, 8188904, 2024. Doi: 10.1155/2024/8188904

[7] H. Alkhalefah, D. Preethi, N. Khare, M. H. Abidi, U. Umer, "Deep learning infused SIRVD model for COVID-19 prediction: XGBoost-SIRVD-LSTM approach," Front. Med. vol. 11, 1427239, 2024. Doi: 10.3389/fmed.2024.1427239.

[8] N. Yalçın and S. Ünaldı, "Symptom Based COVID-19 Prediction Using Machine Learning and Deep Learning Algorithms," Journal of Emerging Computer Technologies, vol. 2, no. 1, pp. 22-29, 2022.

[9] Kaggle, “Symptoms and COVID Presence”, Available online at: https://www.kaggle.com/hemanthhari/symptoms-and-covidpresence (Accessed on Sep. 17, 2024)

[10] A. Tsatsakis, D. Calina, L. Falzone et al., “SARS-CoV-2 pathophysiology and its clinical implications: An integrative overview of the pharmacotherapeutic management of COVID-19,” Food Chem. Toxicol., vol. 146, pp. 111769, 2020.

[11] K. Pearson, “On lines and planes of closest fit to systems of points in space,” Philosophical Magazine, vol.62, no. 1901, pp. 559-572.

[12] S. Wold, K. Esbensen, and P. Geladi, “Principal component analysis,” Chemometrics and Intelligent Laboratory Systems, vol. 2, no. (1-3), pp. 37–52, 1987. doi:10.1016/0169-7439(87)80084-9

[13] L. Peng, G. Han, A. L. Pagou, and J. Shu, “Detecting electric submersible pump failures,” In Machine Learning and Data Science in the Oil and Gas Industry (pp. 223-236), 2021. Gulf Professional Publishing.

[14] Scikit-learn, SelectKBest, Available online at: https://scikit-learn.org/dev/modules/generated/sklearn.feature_selection.SelectKBest.html (Accessed on Oct. 7, 2024)

[15] T. Chen and C. Guestrin, “XGBoost: A Scalable Tree Boosting System, 22nd ACM SIGKDD,” Int. Conf. on Knowledge Discovery and Data Mining, 2016, pp. 785-794.

[16] W. Wang, F. Harrou, B. Bouyeddou, S. M. Senouci, and Y. Sun, “Cyber-Attacks Detection in Industrial Systems Using Artificial Intelligence-Driven Methods,” Int. J. Crit. Infrastruct. Prot., 2022, 38, 100542.

[17] N. Yalçın, S. Çakır, and S. Ünaldı, “Attack Detection Using Artificial Intelligence Methods for SCADA Security,” in IEEE Internet of Things Journal, doi: 10.1109/JIOT.2024.3447876.

Determination of Approximate Values of Spatial Resection in Photogrammetry

Tamás Jancsó
Alba Regia Technical Faculty
Óbuda University
Szekesfehervar, Hungary
jancso.tamas@amk.uni-obuda.hu

***Abstract*— We must have at least four control points for spatial resection when solving with adjustment procedure. The purpose of the spatial resection is to determine the coordinates of the projection center and the three rotation angles that determine the orientation of the image. During the adjustment, we use collinear equations as a basis, which are nonlinear equations with respect to the unknowns. Converting the equations into a Taylor polynomial, we perform a gradual approximation starting from well-chosen starting values. Specifying initial values can cause difficulties if the rotation angles are not close to zero. It can also cause problems when setting the correct initial values if the projection center is not located above the center of the polygon delimiting the control point field. In such cases, the iterative process of the adjustment procedure does not converge to the correct results but diverges or fluctuates around the correct results in an endless cycle, never reaching the error limit necessary to exit the iterative process. In order to choose the appropriate initial values, appropriate initial values can be given to the unknowns by applying the previously developed algorithm without iteration. Testing of the method is presented in MatLab software.**

Keywords*— *photogrammetry, spatial resection, adjustment

I. Introduction

In photogrammetry, the exterior orientation elements (X_O, Y_O, Z_O, φ, ω, κ) of an image are determined, according to the generally accepted method, by spatial resection, with an iteration process [1], see also Fig. 1. and formula (1). Specifying the approximate values of the unknowns is a compulsory task. In some cases, too imprecise starting values can cause divergence or false results in the iteration process. If we solve this task in the case of a pair of stereo images, we have the opportunity to perform a spatial evaluation to produce a DTM (digital terrain model) or a true orthophoto.

To solve the problem, the following steps must be performed:

- Specifying approximate values of unknowns.
- Solving collinear equations according to the Taylor polynomial.
- Writing a system of equations including all measured control points.
- Solving the system of equations by least-square adjustment iteratively.

$$\xi - \xi_0 = -c_k \frac{r_{11}(X-X_O)+r_{21}(Y-Y_O)+r_{31}(Z-Z_O)}{r_{13}(X-X_O)+r_{23}(Y-Y_O)+r_{33}(Z-Z_O)}$$
$$\eta - \eta_0 = -c_k \frac{r_{12}(X-X_O)+r_{22}(Y-Y_O)+r_{32}(Z-Z_O)}{r_{13}(X-X_O)+r_{23}(Y-Y_O)+r_{33}(Z-Z_O)} \quad (1)$$

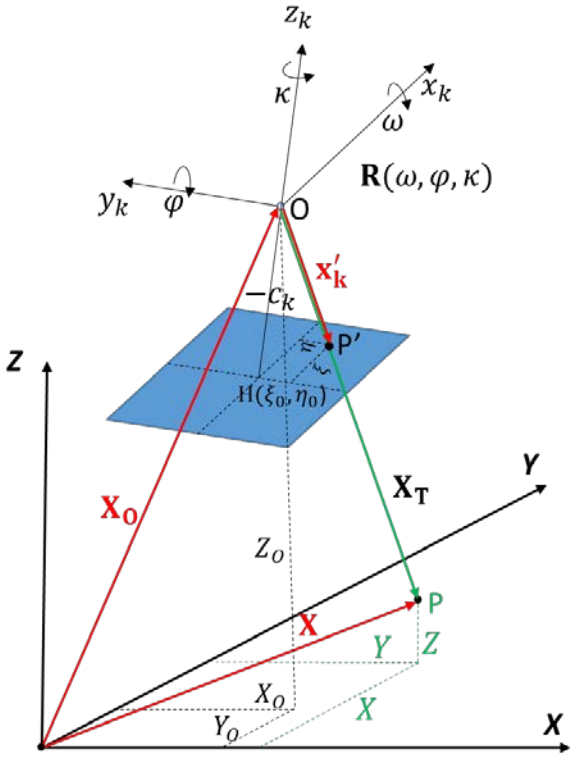


Fig. 1. Spatial resection of one image

II. Specifying Approximate Values

The approximate values required for spatial resection can be specified in the following ways:

- Based on GPS/IMU data.
- Based on flight plan.
- Based on control points.
- Based on the results of an alternative solution.

A. GPS/IMU data as approximate values

GPS coordinates correspond to X_O, Y_O, Z_O values. The rotation angles φ, ω, κ are obtained by conversion from IMU yaw, pitch, roll data. Inertial rotations are relative to geographic north, while photogrammetric rotation angles are relative to the applied projection system. The difference between the two is called meridian convergence, the value of which depends on the coordinate system used.

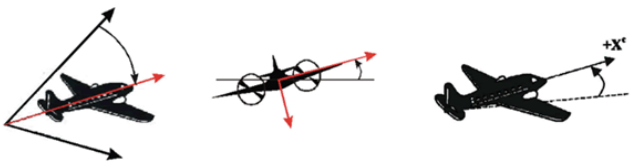


Fig. 2. Yaw, pitch and roll angles interpteted in IMU units

B. *Approximate values based on a flight plans*

The flight plan includes the map locations of projection centers (X_O, Y_O) and the planned flight altitude (Z_O). For the rotation angles, the preliminary value of φ and ω is zero. The angle κ can be specified based on the direction of the planned flight axis.

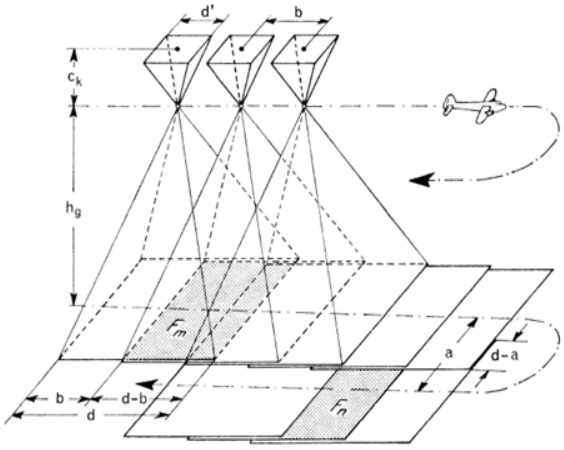


Fig. 3. Main paramaters of a flight plan [8]

C. *Approximate values calculated from the control point coordinates*

The average of the horizontal coordinates of the control points gives the preliminary value of X_O, Y_O. The average of the altitude coordinates will be Z_O , to which we can add the planned or average relative flight altitude h_r. We have no information regarding the rotation angles φ, ω, κ, so we initially take them as zero.

$$X_o = \frac{\sum_{i=1}^{n} X_i}{n}$$
$$Y_o = \frac{\sum_{i=1}^{n} Y_i}{n} \qquad (2)$$
$$Z_o = \frac{\sum_{i=1}^{n} Z_i}{n} + h_r$$

D. *Approximate values based on the results of an alternative solution*

Spatial resection can be solved directly without iterative calculation. There are many examples of this in the literature ([2], [3], [4], [5], [6], [7]). Among them, Jancsó [7], in addition to the solution, also made the error filtering approach part of the solution. The big advantage of this is that the obtained results will be free of gross errors, so the iterative process gets good approximate values, and it will surely converge. The possible ways of using the direct solution are summarized in Fig. 4. The best result can be achieved if we use the method of direct solution with help of all control points in all combinations of four points, see details in [7].

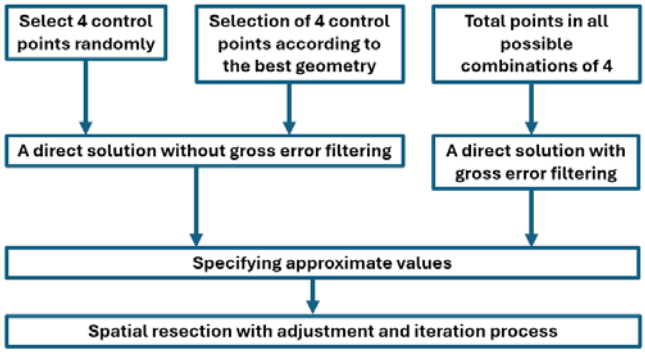


Fig. 4. Possible ways of using an immediate solution

III. Determination of Initial Values for Terrestrial Images

In the case of terrestrial images, we face the following problems when entering initial values:

- The rotation angles can be arbitrary values.
- The location of the projection center can be arbitrary.
- The coordinates X_O , Y_O can even be outside the polygon delimiting the control point field.
- Typically, no GPS/IMU data.

We can see an example of a problematic case in the following chapter.

IV. Calculation Examples

As a first example, we perform the task of spatial resection for the 8 control points measured on an aerial photograph based on the data shown in Table I. The focal length of the camera was 152.7340 mm.

TABLE I. Control Points of an Aerial Photograph

	Ground coordinates			Image coordinates	
No.	**X[m]**	**Y[m]**	**Z[m]**	**ξ[mm]**	**η[mm]**
1	558267.04	6320552.72	44.27	13.63903	110.28369
2	560485.92	6320253.51	5.48	-104.83719	-103.77670
6	557606.30	6317897.56	14.99	6.20823	-0.29857
8	557403.36	6315599.86	15.44	8.12289	88.55405
12	558754.44	6319981.05	24.79	-34.72910	-87.42112
13	558445.39	6318213.81	7.26	-26.24109	-14.89276
14	559312.57	6318041.31	4.80	-61.11469	-10.29072
16	560018.89	6317967.62	7.33	-89.67949	-9.17741

The initial values were calculated according to formula (2) based on the coordinates of the control points, taking into account the approximate flight height as well. The initial values of the rotation angles were taken as zero as it is shown in Table II. The iterative solution is following a MATLAB program created by Bashar Alsadik - University of Twente - ITC in 2010, based on lecture notes of "Analytical photogrammetry" - 2009, Baghdad University - College of engineering - Surveying Department.

TABLE II. INITIAL VALUES BASED ON CONTROL POINTS

Orientation element	Initial value
X_o[m]	558786.7388
Y_o[m]	6318563.43
Z_o[m]	3815.55
φ [deg]	0.0
ω [deg]	0.0
κ [deg]	0.0

After the 6th iteration, we obtained the results summarized in Table III.

TABLE III. RESULT AFTER THE 6TH ITERATION

Orientation element	Adjusted value
X_o[m]	557842.8096
Y_o[m]	6318020.5177
Z_o[m]	3883.4661
φ [deg]	-1.2071
ω [deg]	-2.0722
κ [deg]	-3.6322

In the graph of Fig. 5, we can see the iteration process regarding the refinement of the angles and the camera position. It can be seen from the figure that the changes in the approximate values show a slight fluctuation in iteration 2 and 3.

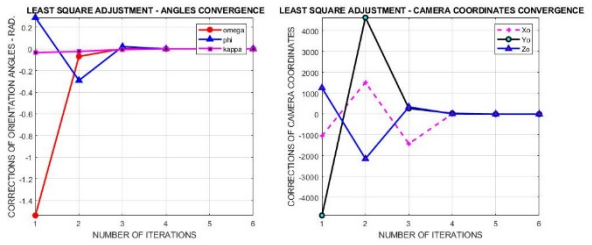


Fig. 5. Iteraton process for an aerial image using all control points

Using the direct method described in [7], we can get better initial values (see Table IV) since we were able to detect the points (point 1, 2 and 6) having gross errors. Thanks to this, the RMS error decreased from the previous value of 0.069 mm to 0.012 mm on the image plane.

TABLE IV. INITIAL VALUES GAINED BY DIRECET METHOD

Orientation element	Initial value
X_o[m]	557846.852
Y_o[m]	6318015.725
Z_o[m]	3886.428
φ [deg]	-1.2762
ω [deg]	-2.0006
κ [deg]	-3.6727

After the3 rd iteration, we obtained the results summarized in Table V.

TABLE V. RESULT AFTER THE 3RD ITERATION

Orientation element	Adjusted value
X_o[m]	557848.1003
Y_o[m]	6318015.1107
Z_o[m]	3886.488
φ [deg]	-1.2948
ω [deg]	-1.9937
κ [deg]	-3.6224

In the graph of Fig. 6, we can see the iteration process regarding the refinement of the angles and the camera position. It can be seen from the figure that the iteration process converges fast without any fluctuation.

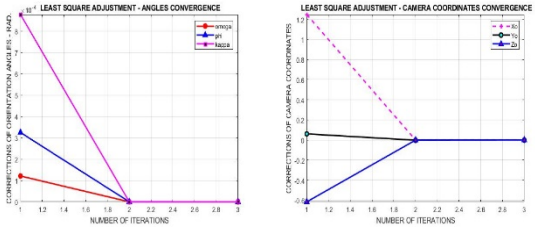


Fig. 6. Iteraton process for an aerial image using control points without gross errors

In Fig. 7 and 8, we can see that if we take the difference between the result of the calculation performed by the direct method and the result of the iteration performed with a good initial value, we can see that hardly any correction was needed during the iteration process.

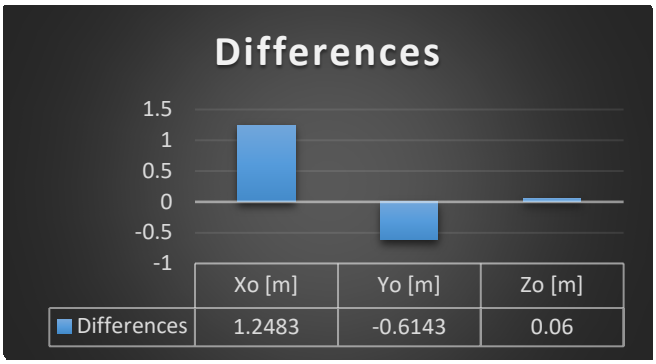


Fig. 7. Differences of rotation angles

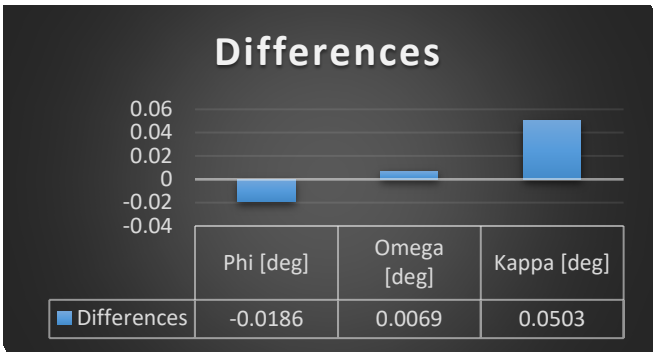


Fig. 8. Differences of camera positions

Finally, let's briefly review an example where the goal was to provide orientation elements of a terrestrial image. Table VI shows the ground and images coordinates. The focal length of the camera was 26.556 mm.

TABLE VI. CONTROL POINTS OF A TERRESTRIAL PHOTOGRAPH

	Ground coordinates			Image coordinates	
No.	**X[m]**	**Y[m]**	**Z[m]**	**ξ[mm]**	**η[mm]**
4	5530.8	3156.7	6576.1	12.472	-3.773
8	6317.8	2926.2	6539.4	12.344	2.280
16	4743.9	2946.0	6553.2	9.080	-7.899
41	6814.7	2416.5	6466.8	7.983	7.573
55	4223.1	2453.9	6485.0	4.772	-9.715
107	7022.4	1629.0	6361.7	-0.834	10.514
123	3982.7	1679.2	6374.3	-0.601	-9.906
175	6790.5	855.51	6262.5	-8.914	8.477
189	4187.4	897.02	6260.5	-5.751	-8.343

In Table VII, the initial values were calculated as an average of coordinates of control points. The rotation angles were set to zero, since we didn't have information about them.

TABLE VII. INITIAL VALUES BASED ON CONTROL POINTS

Orientation element	Initial value
X_o[m]	5512.59
Y_o[m]	2106.67
Z_o[m]	6431.056
φ [deg]	0.0
ω [deg]	0.0
κ [deg]	0.0

In this case, the rotation angles were larger and because of this, the iteration process became lengthy with 22 iterations and uncertain with long fluctuation as it can be seen in Fig. 8. and in Table VIII.

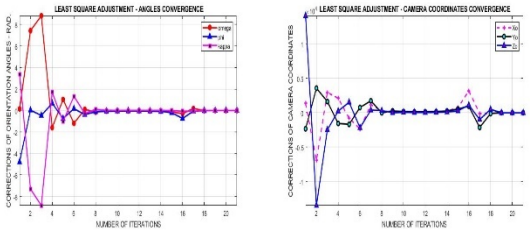


Fig. 9. Iteraton process for a terrestril iamge with large rotation angles

TABLE VIII. RESULT AFTER THE 22ND ITERATION

Orientation element	Adjusted value
X_o[m]	7812.417
Y_o[m]	1694.7481
Z_o[m]	8906.5689
φ [deg]	-397.3723
ω [deg]	810.7068
κ [deg]	-628.3638

After eliminating the points 8, 41 and 175 having gross-errors, we can get better result starting from more accurate initial values calculated by the direct method described in [7]. Fig. 10 shows the smooth iteration process.

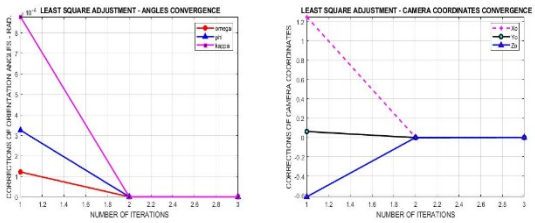


Fig. 10. Iteraton process for a terrestril iamge without gross errors

V. CONCLUSION

In summary, the following important conclusions can be drawn:

- The approximate values of the projection center can be set based on the control points in the case of near-nadir aerial photographs. The initial values of rotation angles are zero.
- For images with a general position, the rotation angles φ, ω, κ can be arbitrary values. In that case, inaccurate approximate values can fail the adjustment procedure.
- It is very important to filter out gross errors when specifying approximate values, otherwise the adjustment accuracy will be low, or it can lead to failure of the adjustment procedure.

REFERENCES

[1] Luhman T., Robson S., Kyle S., Boehm J., "Close-Range Photogrammetry and 3D Imaging", Walter de Gruyter GmbH, Berlin/Boston, 2nd Edition, ISBN 978-3-11-030269-1, 684 p, 2014.

[2] Grunert J.A., "Das pothenotische problem erweiterter gestalt nebst uber seine anwendungen in der geodesie", Grunert' Archive fur Matemathik and Physk, 1:238-248, 1841.

[3] Merrit E.L., "Analytical Photogrametry", Pitman Publishing Co., 242 p., 1958.

[4] Rampal K.K., "A closed solution for space resection", PE&RS, Vol. 45, No. 9, pp 1255-1261, 1979.

[5] Zeng Zh., Wang X., "A general solution of a closed-form space resection", PE&RS, Vol. 58, No. 3, pp 327-338, 1992.

[6] Haralick R.M., "Review and analysis of solutions of the three point perspective pose estimation problem", International Journal of Computer Vision 13(3): pp 331-356, 1994.

[7] Jancsó T., "Gross Error Detection of Control Points with Direct Analytical Method", In: Geo-Imagery Bridging Continents. Proceedings of the XXth ISPRS Congress, Istanbul (IAPRS 35, B3): International Society for Photogrammetry and Remote Sensing (ISPRS), pp 678-682, 2004.

[8] Kraus K.. "Photogrammetry", de Gruyter Textbook, 2nd Edition, pp 459, 2004.

Establishing a Robotic Environment for Reinforcement Learning Study

Bálint Károly Farkas
Alba Regia Faculty, Óbuda University, H-8000
Székesfehérvár, Hungary
farkas.balint@amk.uni-obuda.hu

Károly Széll
Antal Bejczy Center for Intelligent Robotics, Óbuda University
Alba Regia Faculty, Óbuda University, H-8000
Székesfehérvár, Hungary
szell.karoly@amk.uni-obuda.hu

Péter Galambos *Senior Member, IEEE*
Antal Bejczy Center for Intelligent Robotics, Óbuda University
Budapest, Hungary
peter.galambos@irob.uni-obuda.hu

***Abstract*—This paper aims to outline both hardware and software design to create a reliable and efficient framework for future robotics and reinforcement learning research. The primary goal of this paper is to design and establish a functional environment for a reinforcement learning pick-and-place project. This includes the integration of key hardware components, as well as the development of a robust software architecture. The software setup will be built around ROS (Robot Operating System) to provide seamless communication and control of the system. Additionally, plans for software integration will include the use of NVIDIA Isaac Manipulator for AI-driven motion planning and object detection, leveraging tools like cuMotion and FoundationPose. This environment will be a solid foundation to future researches.**

***Index Terms*—Reinforcement Learning; AI-driven motion planning; Autonomous Manipulation; Robotic Research Platform**

I. Introduction

The increasing demand for automation in industries such as manufacturing, assembly and logistics has brought robotic systems to the forefront. Among these, pick-and-place tasks are critical as they require a robot to identify, grasp and move objects from one location to another with precision. This article discusses the setup of a pick and place environment using the UR5e robotic arm, Robotiq gripper and Intel RealSense D435 camera integrated with a NVIDIA GPU-equipped PC running NVIDIA Isaac Manipulator for efficient motion planning and object recognition.

The primary advantage of this project lies in its focus on establishing a modular and adaptable system that can serve as a foundation for advanced research in Reinforcement Learning (RL) for robotic manipulation. By setting up this environment, future research can explore various RL algorithms, allowing the robot to autonomously learn and optimize its behavior in complex tasks, such as picking objects of different shapes, sizes and weights, or dynamically adapting to changes in the environment.

The environment opens up numerous possibilities for advancing RL research in robotics. Researchers can experiment with state-of-the-art RL techniques, like Deep Q-Networks (DQNs), Proximal Policy Optimization (PPO), and Deep Deterministic Policy Gradient (DDPG), to train the robot in simulated or real-world settings [1] [2]. Using the NVIDIA Isaac platform's GPU acceleration, this project can facilitate faster training times and enable the robot to learn intricate tasks through trial and error [3]. It paves the way for future research to enhance robot autonomy.

II. Hardware Setup

Setting up the hardware for a robotic pick-and-place task requires careful selection and integration of various components to ensure efficient and seamless operation. The key elements include the UR5e robotic arm, Robotiq 2F-85 gripper, Intel RealSense D435 camera and a PC. Each component plays a crucial role in the system, providing motion control, object handling, perception and data processing. Ensuring proper connectivity and calibration between these devices is essential to achieve precise and reliable robotic performance.

UR5e Robotic Arm: The UR5e from Universal Robots is a highly versatile 6-axis robotic arm commonly used in research because of its flexibility and the ability to easily integrate with other systems as Toner et al. did [4]. Its compatibility with Robot Operating System (ROS) enables smooth communication with the rest of the system, including the vision system and gripper.

Robotiq Gripper 2F-85: The Robotiq adaptive gripper is a highly adaptable gripper designed to work seamlessly with the UR5e. It supports multiple grasping strategies, including parallel, encompassing and fingertip grasps, allowing it to handle various object shapes and sizes, which is a good solution for us, since it can grip a various type of parts. Its ease of use and quick integration with ROS make it a perfect fit for our setup.

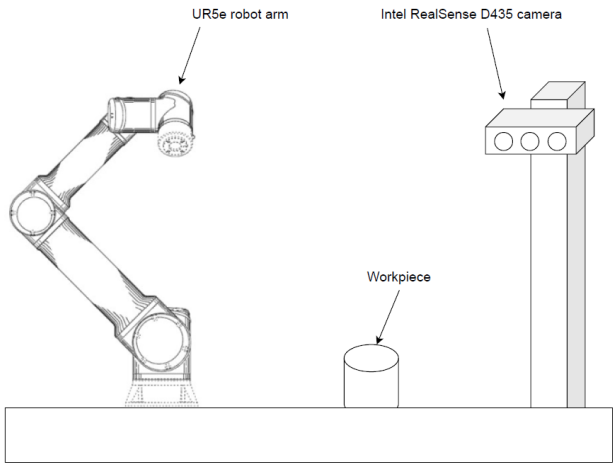


Fig. 1. Spatial arrangement of the components

Intel RealSense D435 Camera: The Intel RealSense D435 camera provides high-quality RGB-D data (color and depth), which is essential for object detection and scene understanding. The camera is capable of real-time depth sensing, making it a valuable tool for detecting objects in 3D space and providing accurate depth information to the robot.

PC: Our PC is equipped with a NVIDIA GeForce RTX 4070 GPU, which is required for the Reinforcement Learning agent and for the visual parts of the research.

After selecting and configuring the core hardware components, it is essential to ensure smooth integration for seamless communication and task execution. Figure 1 shows the spatial arrangement of the system's primary components. The UR5e robotic arm is centrally positioned, with the Robotiq gripper attached to it, communicating with the robotic arms own integrated cable. The Intel RealSense camera is placed at an optimal height and angle to have the best view to the space before the robotic arm.

III. Software Environment

The successful operation of a robotic pick-and-place system depends not only on well integrated hardware, but also on a robust and flexible software environment. This system relies on ROS, an open-source framework that enables seamless communication between hardware components.

In addition, the NVIDIA Isaac Manipulator platform enhances the system by providing powerful AI-driven tools for motion planning, perception and trajectory optimization, leveraging the processing power of an NVIDIA GPU.

The software environment ensures real-time data exchange and synchronization between the various components, allowing the system to perform complex tasks efficiently and accurately. This section outlines the key software tools and libraries involved in integrating the hardware, managing communication, and optimizing task execution.

A. ROS

ROS is a widely-used, open-source framework that provides flexible architecture for developing and integrating robotics software. In this pick-and-place system, ROS serves as the backbone for communication between the different hardware components, such as the UR5e robotic arm, the Robotiq gripper and the Intel RealSense D435 camera.

ROS allows for modular control of the robot through its nodes and topics, where each device of the process is treated as a separate node that communicates through published messages on specific topics.

Furthermore, ROS offers numerous libraries and tools for robotic control, including MoveIt for motion planning and manipulation.

With its open-source nature and extensive community support, ROS ensures flexibility and scalability in robotic applications, enabling the seamless integration of new sensors, actuators, or algorithms as needed.

B. NVIDIA Isaac Manipulator

The NVIDIA Isaac Manipulator is a powerful AI-driven platform built to accelerate the development of robotic systems, specifically focusing on tasks such as manipulation, motion planning, and perception. It integrates seamlessly with ROS and provides a suite of tools and pre-trained models that leverage GPU acceleration to optimize robot performance.

The key components of the Isaac Manipulator include **cuMotion**, a GPU-accelerated motion planner that ensures smooth and efficient trajectory planning, **FoundationPose**, a model for the estimation of the pose of objects in 6D, crucial for pick-and-place tasks.

This platform significantly reduces the development time for robotic applications by providing ready-to-use workflows and AI models, making it easier to implement complex tasks while optimizing performance through parallel processing on GPU.

C. Object Detection and Perception

In a pick-and-place robotic system, object detection and perception are crucial to allow the robot to interact with its environment effectively. Advanced object detection algorithms like SyntheticaDETR provide real-time object detection capabilities, enabling precise and adaptive manipulation. However, the robustness of perception systems is often challenged by external factors and software-related issues that may cause errors. Implementing fault-tolerant solutions, similar to those used in microcontroller-based systems, can help mitigate these risks [5]. Our camera is a RGB-D camera, which means it gives us an RGB picture plus depth information. This depth information is curcial for accurate object detection.

To interpret the camera data, the system utilizes advanced object detection algorithms like **SyntheticaDETR** for NVIDIA Isaac, which is designed for real-time object detection in complex environments.

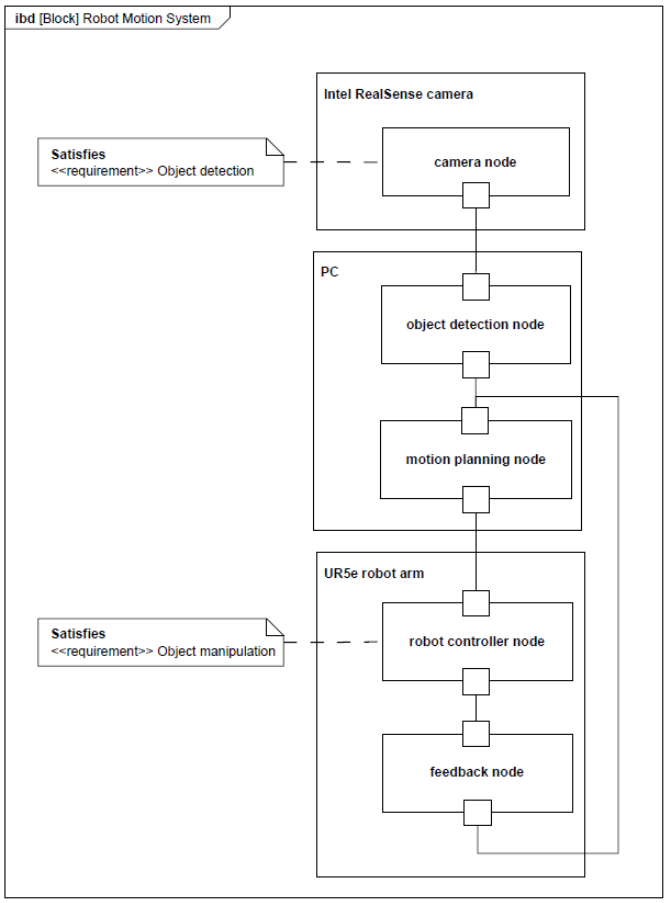


Fig. 2. Communication pathways

This perception tool allows the robot to dynamically adjust its actions based on real-time data ensuring precise grasping and manipulation of objects.

The overall software architecture of the system, including its communication pathways and data flows, is depicted in figure 2, providing a clear visualization of how the components are interconnected and how data are processed throughout the workflow.

D. Implementing Software Environment

The software integration plan involves setting up and configuring the software components needed for communication, perception, and control of the robotic system.

ROS2 Jammy: ROS is installed on the main PC and the Raspberry Pi 5, with necessary ROS packages for the UR5e, Robotiq gripper, and RealSense D435 camera. This setup ensures that each hardware component is treated as a ROS node capable of publishing and subscribing to topics for data exchange. Nodes for motion control, gripper operation, and sensor data processing are launched to create a cohesive system.

NVIDIA Isaac Integration: The NVIDIA Isaac Manipulator software stack is installed on the PC equipped with an NVIDIA GPU. The integration process involves setting up the cuMotion module for motion planning and the FoundationPose model for object detection. These components are integrated into the ROS ecosystem, allowing them to interact with other ROS nodes seamlessly. Isaac's perception tools are configured to process the depth and RGB data from the RealSense camera to identify and localize objects in the workspace.

To set up the software environment, I followed these steps:

1) **Windows Subsystem for Linux (WSL) Installation**
2) **NVIDIA Drivers Installation**
3) **CUDA Installation**
4) **Docker Setup**
5) **NVIDIA Container Toolkit Installation**
6) **Creating an Isaac ROS Workspace**
7) **Cloning Isaac ROS Common**
8) **Cloning RealSense ROS**
9) **Cloning Isaac Manipulator**
10) **Running "run_dev.sh" Script**

IV. Future Work and Research Opportunities

The setup and design of this pick-and-place robotic environment open up a wide array of future research directions, particularly in the field of RL and autonomous robotic manipulation. By establishing a robust hardware and software foundation, this system can be used to explore and develop advanced algorithms that push the boundaries of current robotic capabilities.

A. RL for Adaptive Manipulation

One of the primary opportunities for future work involves implementing and experimenting with various RL algorithms to enable adaptive and intelligent manipulation. With the existing environment, researchers can train the UR5e robotic arm to autonomously learn optimal grasping strategies for objects of varying shapes, sizes and materials. Advanced algorithms such as PPO and DQN can be utilized to enable the robot to adapt its behavior dynamically in response to changes in the environment, such as moving targets or unpredictable obstacles [6]. This research can lead to the development of robots capable of performing complex tasks in unstructured environments, such as logistics, e-commerce warehouses, medical laboratories.

B. Transfer Learning and Sim-to-Real Transfer

Another potential area of research is the use of transfer learning and sim-to-real transfer techniques. By leveraging the NVIDIA Isaac Simulator in conjunction with the real-world environment, the robot can be trained in a simulated environment where it can learn efficiently without the risk of hardware damage [7]. The learned policies can then be transferred to the real robot, minimizing the sim-to-real gap. This approach can significantly reduce the time and resources required for training complex behaviors, providing a more efficient pathway for deploying robots in real-world scenarios [8].

C. Advanced Perception and Object Recognition

Future work can also focus on enhancing the perception capabilities of the system. By integrating more sophisticated object detection algorithms, such as Convolutional Neural Networks (CNN), the robot can be trained to recognize and manipulate a wider variety of objects with greater accuracy. Additionally, incorporating scene understanding and semantic segmentation could allow the robot to distinguish between different types of object and make context-aware decisions during pick-and-place tasks.

D. Collaborative and Multi-Robot Systems

The current setup provides a single-robot environment, but future research could expand this to collaborative or multi-robot systems. By integrating additional robotic arms or mobile robots into the environment, researchers can explore how multiple agents can work together to accomplish complex tasks, such as cooperative assembly, sorting or task allocation [9].

E. Real-Time Adaptation and Fault Tolerance

Another potential area of exploration is the development of algorithms for real-time adaptation and fault tolerance. In real-world scenarios, robots must be capable of handling unexpected situations, such as hardware malfunctions, sensor noise, or changes in object properties. Implementing fault detection and recovery strategies, combined with RL-based adaptive learning, can enable the robot to continue operating effectively even in the face of unforeseen challenges, increasing the robustness and reliability [10].

By exploring these avenues, the current environment can serve as a versatile platform for advancing the state-of-the-art in robotic manipulation, RL. Each research direction offers the potential to significantly enhance the capabilities of robotic systems, driving progress in both academic research and practical industrial applications.

V. Conclusion

This project focused on the comprehensive design and setup of an environment tailored for a robotic pick and place system using the UR5e robotic arm, Robotiq gripper and Intel RealSense D435 camera. The main objective was to establish a solid foundation for the system, ensuring that all hardware and software components were carefully selected, configured, and integrated to facilitate future development and implementation of pick and place tasks.

Throughout the project, critical aspects such as hardware positioning, software integration through ROS, and the incorporation of NVIDIA Isaac Manipulator tools were addressed to ensure seamless communication and data processing within the system. Special attention was paid to the calibration and configuration of each component to establish a synchronized environment capable of supporting complex robotic operations.

Although the system is not yet fully operational, the groundwork laid during this project is crucial for future developments. By thoroughly designing and setting up the environment, we have created a robust platform that will allow further exploration, testing, and optimization of pick-and-place algorithms. This environment serves as a foundational step towards a fully functional robotic system, allowing future work to focus on refining task execution, improving efficiency, and addressing real-world automation challenges.

Acknowledgment

Péter Galambos and Károly Széll thankfully acknowledge the financial support of this work by project no. 2019-1.3.1-KK-2019-00007 implemented with the support provided from the National Research, Development and Innovation Fund of Hungary, financed under the 2019-1.3.1-KK funding scheme. Péter Galambos is a Bolyai Fellow of the Hungarian Academy of Sciences. Péter Galambos is supported by the UNKP-22-5 (Bolyai+) New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund.

Bálint Farkas is supported by the 2024-2.1.1 university research scholarship program of the ministry for culture and innovation from the source of the national research, development and innovation fund.

The authors acknowledge the use of ChatGPT, a language model developed by OpenAI, in an assistive and supervised manner for generating and refining content in this article. ChatGPT was utilized to draft and enhance sections such as the Abstract, Introduction, and Conclusion. The AI-generated content was carefully reviewed and edited by the authors to ensure accuracy and alignment with the research objectives.

References

[1] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "Deep Reinforcement Learning: A Brief Survey," *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 26–38, 2017. [Online]. Available: https://ieeexplore.ieee.org/document/8103164

[2] M. Q. Mohammed, K. L. Chung, and C. S. Chyi, "Review of Deep Reinforcement Learning-Based Object Grasping: Techniques, Open Challenges, and Recommendations," *IEEE Access*, vol. 8, pp. 178 450–178 481, 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9210095

[3] A. Goyal, V. Blukis, J. Xu, Y. Guo, Y.-W. Chao, and D. Fox, "RVT-2: Learning Precise Manipulation from Few Demonstrations," Jun. 2024, arXiv:2406.08545 [cs]. [Online]. Available: http://arxiv.org/abs/2406.08545

[4] T. Toner, M. Saez, D. M. Tilbury, and K. Barton, "Opportunities and challenges in applying reinforcement learning to robotic manipulation: An industrial case study," *Manufacturing Letters*, vol. 35, pp. 1019–1030, Aug. 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2213846323001128

[5] G. Györök and B. Beszédes, "Fault-tolerant Software Solutions in Microcontroller Based Systems," 2017.

[6] Z. Zhou, R. Zuo, B. Ying, J. Zhu, Y. Wang, X. Wang, and X. Liu, "A Sensory Soft Robotic Gripper Capable of Learning-Based Object Recognition and Force-Controlled Grasping," *IEEE Transactions on Automation Science and Engineering*, vol. 21, no. 1, pp. 844–854, 2024, conference Name: IEEE Transactions on Automation Science and Engineering. [Online]. Available: https://ieeexplore.ieee.org/document/9999176

[7] "Closing the Sim-to-Real Gap: Training Spot Quadruped Locomotion with NVIDIA Isaac Lab," 2024, publication Title: NVIDIA Technical Blog. [Online]. Available: https://developer.nvidia.com/blog/closing-the-sim-to-real-gap-training-spot-quadruped-locomotion-with-nvidia-isaac-lab/

[8] D. Horváth, K. Bocsi, G. Erdős, and Z. Istenes, "Sim2Real Grasp Pose Estimation for Adaptive Robotic Applications," *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 5233–5239, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2405896323004676

[9] X. Lan, Y. Qiao, and B. Lee, "Application of Multi Agent Reinforcement Learning to Robotic Pick and Place System," in *2024 10th International Conference on Automation, Robotics and Applications (ICARA)*, 2024, pp. 1–6. [Online]. Available: https://ieeexplore.ieee.org/document/10552987

[10] H. Xue, C. Pan, Z. Yi, G. Qu, and G. Shi, "Full-Order Sampling-Based MPC for Torque-Level Locomotion Control via Diffusion-Style Annealing," Sep. 2024, arXiv:2409.15610 [cs]. [Online]. Available: http://arxiv.org/abs/2409.15610

Colour based LEGO robot controlling

Péter Udvardy
Alba Regia Faculty
Institute of Science and Software Technology
Obuda University
Szekesfehervar, Hungary
udvardy.peter@amk.uni-obuda.hu

György András Györök
Alba Regia Faculty
Obuda University
Szekesfehervar, Hungary
gyorok.andras@uni-obuda.hu

Bertalan Beszédes
Alba Regia Faculty
Institute of Engineering
Szekesfehervar, Hungary
beszedes.bertalan@amk.uni-obuda.hu

***Abstract*— Collaborative robots are capable of work performed jointly by humans and robots, the safety aspects of which are of fundamental importance in relation to the design of the work environment. Robots and people can only move freely in the workplace in appropriate places, there are also prohibited areas for robots and workers. To differentiate this, the colour-based work area designation is one of the simplest and, in some respects, cost-effective solutions. During our tests, we used the LEGO Mindstorms EV3 robot to examine the detection capabilities of the robot's colour and light sensor, the effect of different programming options on the robot's movement, and the proper detection of the red and blue areas. As a summary, we can draw the conclusion that the sensors of mobile robots are suitable for navigation purposes, they are able to accurately recognize and distinguish colours and can therefore be used well even in an industrial environment to achieve both navigation and related occupational safety goals.**

Keywords*— *LEGO, robot, color based control

I. LEGO ROBOTICS

With the development of robotics, there is an increasing demand for the environmental interaction of an intelligent device to be as large as possible. This is especially true for collaborative robotic techniques when the robot has to work with a real person, an operator and a workspace.

Environmental information does not only mean the exact definition of an object, but in many cases the image and colour information obtained about the object is what is formulated as an input parameter. Nowadays, robot control also assumes as many input parameters as possible. The careful design of the integration of robots into the environment is necessary, it raises not only control but also safety issues.

Nearly 60% of the human brain contains image data, and a certain part contains colour information. If we want to manage the robot in a human-compliant manner, it would be very beneficial to implement the management of colour information as input parameters. Such an input parameter can not only support the robot's own "decision", but can also be a control parameter that can even decide the robot's route, the robot's movement, the creation of trajectories, the exact spatial coordinate of each other. Image processing with a single camera system is obviously suitable for this, the use of much simpler colour sensors, which can produce a well-defined spectral dependence, can give significant results.

The Danish LEGO -obviously due to keeping the market used has changed with the times and offers a component and building element environment that is more than a game, but a real development environment. It is worth getting to know LEGO robots at a very young age, so preschoolers can already meet the concept of algorithm, robot, program, or decision tree. The sensors of the robots are properly programmed, we can give our robot countless instructions, and with the help of these we can even simulate artificial intelligence and free will, it is up to us to equip our machine with what features. These qualities can simulate for example hyperactivity, disinterest, joy, kindness, injury. The hyperactive robot moves forward without stopping until it detects an obstacle on the first sensor (it can be anything, a tree, a table, a wall, or even another robot that is currently in motion), here a random generator draws a lottery that the robot turn right or left, then continue on your way, ad infinitum [1].

II. RGB TECHNICS AND COLOUR SENSORS

RBG is derived from the English terms red, green, blue and supports the creation of different shades of colors based on an additive model. This is based on the fact that any color can be produced from a mixture of red, green and blue colors of different strengths. Based on this, all colors consist of these three main components. The intensity of the individual basic components (RGB) is changed between 0 and 255 values in the case of eight-bit representation. Depending on the desired color, if all three components are set to the maximum value, a white color is obtained. (255R, 255G, 255B) By definition, the code for the pure red color; 255R, 0G, 0B.

Figure 1 illustrates the resolution of the visible range into three basic colors. To mix a "specific color" in the figure; 0.08R+0.96G+0.02B component required. [2]

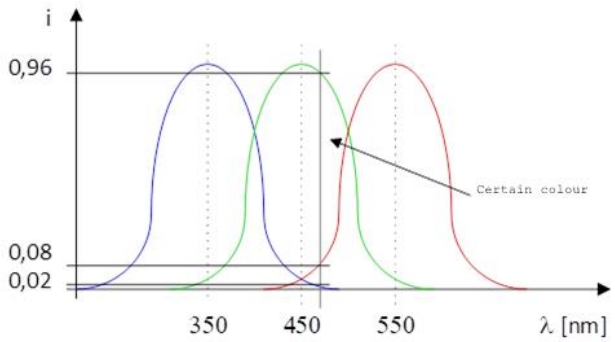


Fig. 1. Additive RGB technics [2]

The RBG color model is used in many places, televisions, monitors, digital displays, where all colors can be displayed by combining the three colors. The two-way interpretation of the model is self-explanatory, since during imaging and color measurement, the R, G, B components of the source of the given wavelength must be determined as coefficients.

Among the light sources, the most well-known light source that operates our civilization is the sun, which provides heat, light and energy to sustain life on earth. The light produced by the chemical reaction of individual organisms is called bioluminescence. Thus, in the case of individual insects or deep-sea animals, these characteristics typically serve the survival of the given individual, may play a role in hiding, hunting, communication, or even self-reproduction. Incandescent lamps and halogen lighting fixtures are traditionally the most common artificial light sources, and they can be of various types and shapes. These lighting devices work with electricity and produce light with very poor efficiency. A few percent of the absorbed electrical power will be visible light, the majority product will be infrared light and heat.

Lasers are strong, intense and uniform light sources that can be widely used in industrial devices, communication, telecommunications, measuring technology, medical applications, measurements, or even in the entertainment industry. Nowadays, the most common ones are based on semiconductors, Laser LEDs (LLED). their spectrum is very small (monochromes). Xenon light sources, in addition to everyday lighting, they are used in decorations and illuminated signs. The light source is xenon gas excited by electric current and driven by a current generator. Their efficiency and life expectancy is much better than that of incandescent lamps.

An LED is a semiconductor-based light source that uses electroluminescence to emit light. When an electric current is applied to the semiconductor, it emits photons during the processes taking place at the P-N junction. We can distinguish several types, even depending on the use. SMD LEDs, due to their size, they are mainly used in places where the possibilities are limited, in terms of size, for example, they are used on printed circuit boards or displays. COB LEDs are high-performance light sources that are used in lamps, which can be achieved by using several discrete LEDs. RGB LEDs combine the three basic colors, screen technology, background light, production of variable color temperature.

The UV LED emits radiation in the ultraviolet spectrum, which is used for medical purposes, sterilization, or the activation of fluorescent materials. The IR LED provides infrared radiation and is used in night vision cameras and remote controls. The infrared range is also often used because Si-based semiconductors work with the best efficiency at this wavelength (780-900nm, depending on additives). High Power LED is mainly used for lighting in cars, public areas, and households. LEDs are well-deserved light sources, as their lifetime, efficiency, and parameter constancy significantly exceed those of the above-mentioned devices. The automation of production has been solved, so the relative costs are also more favorable.

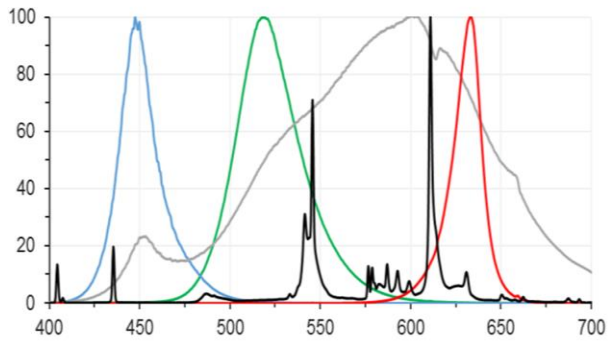


Fig. 2. Spectral intensity of different LED souces compared to a composite light source [3]

The light sensor converts the intensity of the light in its environment into an electrical voltage signal. During operation, the energy of electrons stimulated by incoming photons exceeds the threshold energy of the depleted layer of a silicon semiconductor PN junction. As a result, the UAK voltage of the PN junction increases. Ideally, this is the photodiode, or in a larger size, depending on the technology, the solar cell. It can be said that all semiconductors are photosensitive, since the PN transition depends on the amount of free electrons everywhere. Color perception means determining the RGB coefficients. One way to do this is to choose a semiconductor that is sensitive to the appropriate wavelength. [4] [5]

The other more frequently used method is when choosing a broad spectrum sensor. In front of each photodetector that detects the three basic colors, I place the red, green and blue basic color filter. To "separate" the components of the three basic colors, the color separation mirror process is also used, primarily in imaging devices (CCD) and high-resolution cameras.

The light sensor of LEGO Mindstorms EV3 offers several functional possibilities, depending on the perception of colors, in the navigation of the robot case. The sensor can be used in two ways, one is the reflectance number, which detects the reflected strength of natural or artificial light in the environment. It is able to recognize, avoid and move objects, it recognizes lines even based on color, so we can navigate our robot depending on the program. [6]

The other option is the light source number, which detects the light sources directed at it. This can also work as described above, and allows absolute coordinate navigation, object detection, and distance determination. In the case of LEGO Mindstorms EV3, the setting of the sensors can strongly influence the result obtained. It does not matter at all what angle the incident light receives the sensor, which can be adjusted with the correct positioning to achieve the desired result. [7] [8]

Light-based robot navigation can work in two ways, LIDAR (light detection and ranging) and optical flow. In the case of LIDAR technology, the system emits laser beams, then detects the beams reflected from the objects, and based on the measurement of the reflection time, the system creates a three-dimensional map of the environment. LiDAR is particularly useful in complex, dynamic environments because it can determine the distance and position of objects with high accuracy. It was first used to survey larger areas, as it 'sees' into the subsurface layers. In the case of optical flow, the robot

uses cameras to process visual information from the environment.

Optical flow is based on the analysis of movement patterns between successive frames, which allows the robot to determine speed and direction. This technique is particularly beneficial in indoor environments where the GPS signal is weak or unavailable. It is important to mention that the visual displacement is not necessarily the same as the real displacement.

The use of autonomous robots in healthcare applications can offer significant benefits in hospitals and other healthcare facilities. With the help of light-based navigation, robots can safely and efficiently transport medicines, medical devices and other necessary items between different departments. This reduces the potential for human error and increases operational efficiency. This technology is used in many industrial production plants (for example, pharmaceutical factories). Despite how widespread these solutions are and light-based robot navigation offers many advantages, it also faces many challenges. One of the main challenges is cost-effectiveness, as LIDAR systems can be expensive, which limits the widespread adoption of the technology. In addition, light-based navigation systems can be sensitive to environmental factors, such as changes in light conditions or dust, which can affect the accuracy of the system. In industrial conditions, this is a big challenge for operators. [9] [10]

III. Modelling and results

In order to create the most realistic artificial industrial environment model possible, the system was assembled over a uniform, smooth, white surface, the PAR RGBW led light sources illuminated the surface from above. Three pieces of PAR RGBW led were used, with which the lighting of the model environment could be divided into three sections and the lighting of these sections could be changed in small steps.

The selection of the surface was based on a smooth surface so that the LEGO Mindstorms EV3 robot could easily move and turn on it if necessary, as well as a non-reflective surface, which ensured the correct detection of the color and light sensor. [11] [12]

The PAR RGBW LED lamps used for lighting are equipped with 54 3-watt LED light sources, the red, green, blue and white LEDs are located separately, so they can be regulated separately (12-R, 18-G, 18-B, 6-W). The total power of these light sources is 162 watts. PAR lights can have a variety of functions, such as built-in programs, DMX control or wireless remote control, allowing for ease of use and versatile setup. These lamps are often used in live stage performances, concerts, DJ sets, clubs, but they can also serve as a popular lighting device in the home environment, for example as mood lighting or to create light effects. In addition to hobby use, they are also suitable for creating professional applications.

A DMX connector is available for controlling the LED light sources, where the light intensity and, of course, other properties (such as the flashing and its frequency) can be infinitely adjusted between 0 and 100 values. The DMX control of the PAR RGBW led light source (Figure 8-5), the color composition, intensity and possibly flashing of the light was solved with an FTS DMX 512 controller. DMX 512 is a standard communication protocol in lighting that enables communication and control of lighting devices such as LED lights, dimmers, LED moving light heads and effects. "DMX" stands for "Digital Multiplex", while "512" indicates the transmission capacity of the protocol, which means 512 channels.

This protocol works through a serial digital communication system that provides one-way data transfer. Each DMX device has a unique address (DMX address) that can be specified by setting the control unit. The control unit sends the DMX signal messages, which contain the parameters of the various lighting effects, such as color, brightness or movement. The DMX 512 protocol is very flexible and allows the simultaneous control of many devices. It is often used in live stage performances and concerts, where precise and synchronized lighting effects are required. The protocol enables the creation of synchronized light effects, as well as the dynamic control and programming of lights.

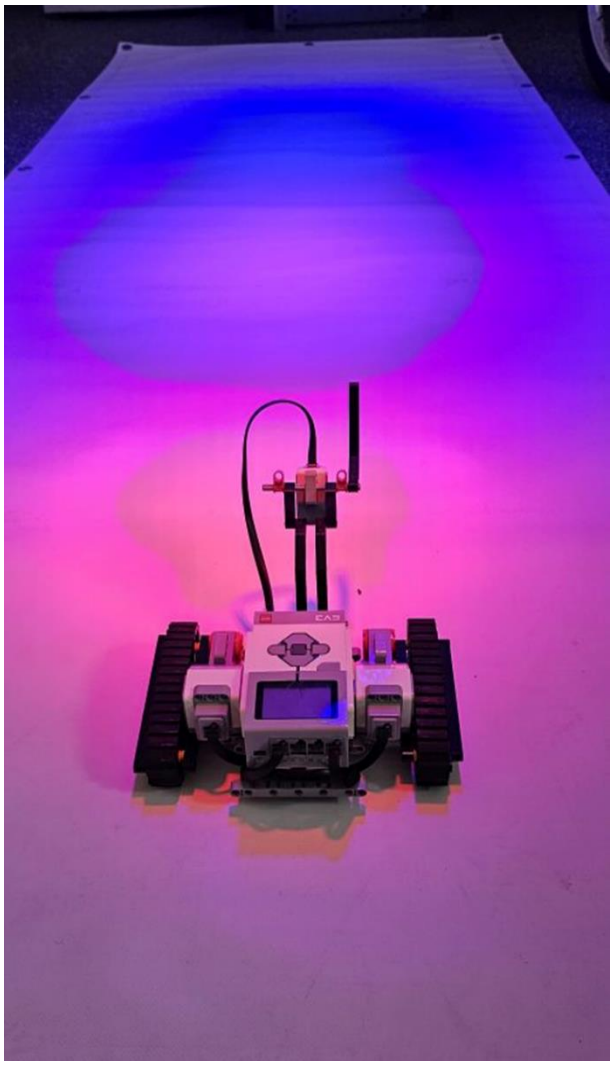


Fig. 3. Working area with the lighting [1]

We also programmed the LEGO Mindstorms EV3 robot on its own programming interface, where the robot, simulating industrial conditions (Figures 8-10, 8-11), moves forward or backward between sectors marked with different colors according to whether it is red, blue or both is the surface illuminated.

In the blue-lit areas, the mobile robot can move freely, which in this case means the forward movement of the EV3 robot with tank drive, then the crawler on both sides rotates at 75% of the maximum engine speed. The areas illuminated by red light are not accessible to the robot, only the worker can move freely in this zone. The EV3 robot detects red color and starts moving backwards, its movement is in a straight line due to the value of -5 on both sides of the tank drive. After leaving the forbidden area, the robot stops and starts again in the other direction. If the space is illuminated with both lights, both the robot and the worker can use it freely. In this case, the EV3 robot continues to move.

Fig. 4. LED light sources [1]

The lighting of the work area is uniform due to the location of the three PAR LED illuminators, the light of the three illuminators overlaps a little, and the colors can be changed freely. In the case of our current model, a part illuminated with both colors was located between the red and blue areas, so the robot changed its movement at the border areas. The color fidelity of the above images is not perfect, in reality the red, blue and the area containing both are well recognized even by the robot.

Through the MATLAB interface, we tested what the robot "sees" during its progress. We took into account the change in the intensity of the red and blue colors, where the robot starts to perceive the second color and how long it perceives the first. In the first step, we checked the connection between LEGO Mindstorms EV3 and the computer. As a second step, an empty graph was created, which is constantly updated with new measured data. The Y-axis of the graph shows the intensity values (0-100) and we read the color sensor continuously for 60 seconds using a while loop.

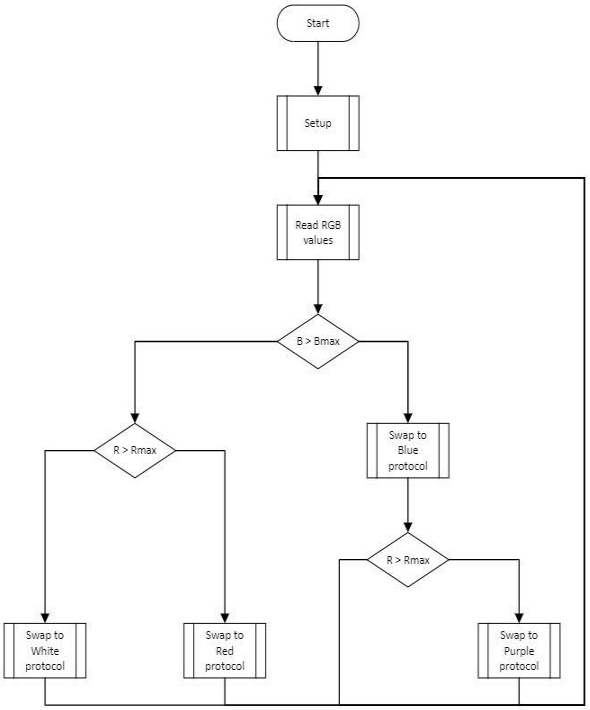


Fig. 5. Programme flowchart [1]

The intensity of the red and then the blue color was read and displayed on the graph. During data display, the graph was updated in real time with the new data using the animated line and draw-now functions. This code ensures that the intensity of the red color measured by the LEGO EV3 color sensor is continuously displayed on a graph in MATLAB, so it is easy to follow the change in the intensity of the red color over time. In the case of the color blue, we proceeded in a similar way. [10] [13]

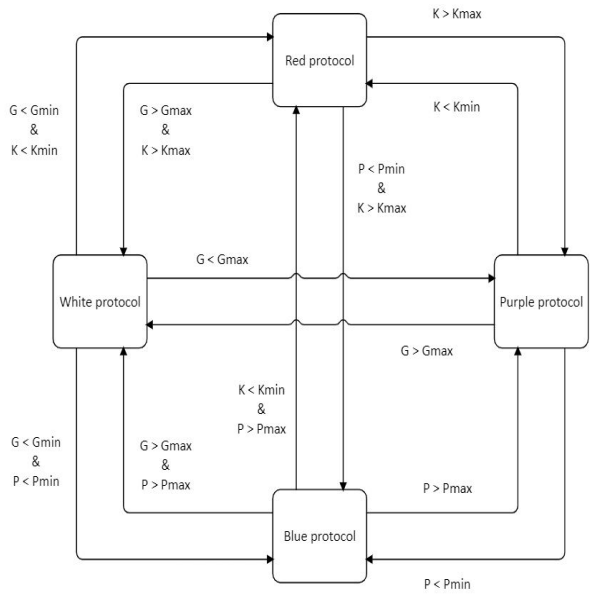


Fig. 6. Status diagram [1]

It can be concluded that both the red and blue colors were well recognized by the EV3 robot, during its forward progress the location of the color transition can be clearly recognized on the graph, the decreasing intensity of red, while the increasing intensity of blue can be detected. This can be significant because, in industrial conditions, the robot, like a human, can distinguish the areas marked with light and change its movement speed and direction accordingly.

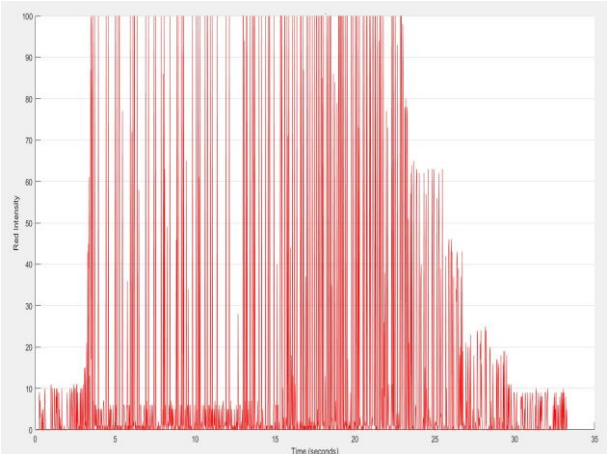


Fig. 7. Changes in red colour intensity [10]

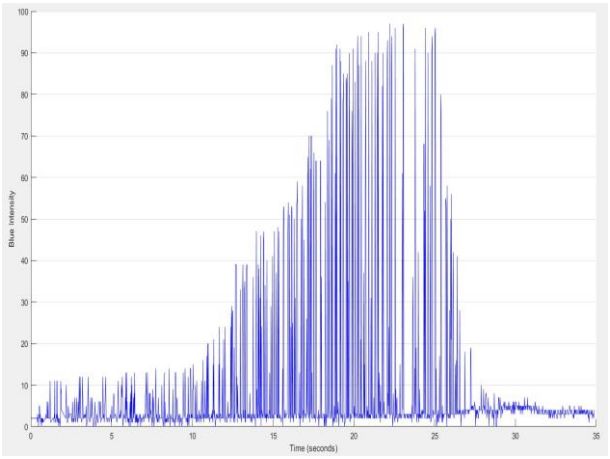


Fig. 8. Changes in blue colour intensity [10]

IV. Summary

One of the conditions for the use of mobile robots in industrial production is that they can find their way around their environment independently, which they can solve in several ways. One way of doing this is light and color perception, when they navigate by recognizing them during their movements. With the help of the sensors of the LEGO Mindstorms EV3 robot, light-based mobile robot navigation can be well simulated, and the behavior of the robots can be modeled by creating the appropriate industrial environment.

Collaborative robots are capable of work performed jointly by humans and robots, the safety aspects of which are of fundamental importance in relation to the design of the work environment. Robots and people can only move freely in the workplace in appropriate places, there are also prohibited areas for robots and workers. To differentiate this, color-based work area designation is one of the simplest and, in some respects, cost-effective solutions.

During our tests, we used the LEGO Mindstorms EV3 robot to examine the detection capabilities of the robot's color and light sensor, the effect of different programming options on the robot's movement, and the proper detection of the red and blue areas.

As a summary, we can draw the conclusion that the sensors of mobile robots are suitable for navigation purposes, they are able to accurately recognize and distinguish colors and can therefore be used well even in an industrial environment to achieve both navigation and related occupational safety goals. The concept of color-marked work areas is suitable for the widespread use of collaborative robots.

References

[1] https://www.lego.com

[2] Gy. Györök: Perifériák OE AREK 8003, Budapest 2013

[3] A. J. von Wangelin: A flow reactor setup for photochemistry of biphasic gas/liquid reactions, DOI: 10.3762/bjoc.12.170

[4] L. Pérez, i. Rodríguez, N. Rodríguez, R. Usamentiaga, and D. F. García, “Robot Guidance Using Machine Vision Techniques in Industrial Environments: A Comparative Review,” Sensors, vol. 16, p. 335, Mar. 2016.

[5] Gy. Györök, B. Beszédes: Artificial Education Process Environment for Embedded Systems. In: Orosz, Gábor Tamás (szerk.) 9th International Symposium on Applied Informatics and Related Areas - AIS2014.Székesfehérvár, Magyarország : Óbudai Egyetem (2014) pp. 37-42. , 6 p.

[6] B. Altakhayneh: The Impact of Using the LEGO Education Program on Mathematics Achievement of Different Levels of Elementary Students. European Journal of Educational Research, v9 n2 p603-610 2020

[7] https://uk.mathworks.com/hardware-support/lego-mindstorms-ev3-matlab.html

[8] https://uk.mathworks.com/hardware-support/lego-mindstorms-ev3-simulink.html

[9] M. Seebauer, P. Udvardy: Collective Behavior Simulation of Mobile Robots Using Lego NXT Sets In: Orosz, Gábor (szerk.) AIS 2018 - 13th International Symposium on Applied Informatics and Related Areas Székesfehérvár, Magyarország : Óbudai Egyetem, Alba Regia Műszaki Kar (2018) pp. 122-125. , 4 p.

[10] https://matworks.com

[11] N. Montes et al.: A novel educational platform based on matlab/simulink/lego ev3 for teaching with robots, 2018. INTED2018 proceedings, pp. 975-980. ISSN: 2340-1079

[12] P. Udvardy: LEGO robotics at service of education In: Petőné Csuka Ildikó (szerk.) AIS 2022 - 17th International Symposium on Applied Informatics and Related Areas, Székesfehérvár: Óbudai Egyetem, pp 33-37 (2022)

[13] P. Udvardy et al.: 3D modelling by UAV survey in a church,2019 New Trends in Aviation Development (NTAD),189-192,(2019)

Data Integration between Mixed Reality Software and Industrial Robot Controller via OPC UA

Zoltán Szilágyi
Doctoral School of Applied Informatics and Applied Mathematics, Óbuda University
Budapest, Hungary
szilagyi.zoltan@amk.uni-obuda.hu

Éva Hajnal
Alba Regia Faculty, Óbuda University
Székesfehérvár, Hungary
hajnal.eva@amk.uni-obuda.hu

Imre Csóka
Alba Regia Faculty, Óbuda University
Székesfehérvár, Hungary
csokaimre@stud.uni-obuda.hu

Péter Béres
Alba Regia Faculty, Óbuda University
Székesfehérvár, Hungary
beres0610@gmail.com

Károly Széll
Antal Bejczy Center for Intelligent Robotics Alba Regia Faculty, Óbuda University
Székesfehérvár, Hungary
szell.karoly@amk.uni-obuda.hu

Péter Galambos
Antal Bejczy Center for Intelligent Robotics, Óbuda University
Budapest, Hungary
peter.galambos@irob.uni-obuda.hu

***Abstract*—Industrial robot production lines require the development of advanced communication protocols that can adapt to the diverse needs of the interconnected devices. In this article, we present a solution that employs the sophisticated industrial communication capabilities of robots to provide data for an augmented reality system designed to enhance the robots' capabilities. The solution is based on OPC UA, a unified, secure, and reliable communication framework between industrial devices and software. A Python-based OPC client installed on a Raspberry Pi device can successfully read data from an industrial robot via an OPC-UA connection. The test results demonstrated the viability of the solution and provided the opportunity to develop the availability of all the data to be transferred to HoloLens.**

***Index Terms*—mixed reality, industrial robotics, OPC UA, JSON**

The OE-RH-423/2024 project has been implemented with the support of doctoral research work carried out at company-university cooperation - relevant to domestic R&D programs – provided by "Óbuda University Cooperative PhD Student Scholarship Program", from the "Talent Care Fund" of Óbuda University. Péter Galambos is a Bolyai Fellow of the Hungarian Academy of Sciences. Péter Galambos is supported by the UNKP-22-5 (Bolyai+) New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund.

I. Introduction

The increasing integration of industrial robot production lines necessitates the development of advanced communication protocols that can adapt to the diverse needs of the interconnected devices. The solution presented in this article employs the sophisticated industrial communication capabilities of robots to provide data for an augmented reality system designed to enhance the robots' capabilities. A plethora of industrial communication solutions exists, optimized for a multitude of tasks. Such protocols include a variety of industrial bus or Ethernet protocols, MQTT [1], [2], and numerous variants of OPC. From the plethora of available communication options, it was imperative to select a solution that would satisfy the following criteria:

- This communication method is a widely adopted standard within the industry. A significant number of the target devices are capable of establishing connections via the selected communication method.
- It is imperative that the selected communication method be secure. It fulfills both industrial and task-specific security requirements.
- It provides access to a plethora of data.
- It enables bidirectional communication.

The other aspect of the selection criteria was the data

receiving and processing unit, which, being a non-industrial device, typically supports different communication forms. A communication solution had to be selected that is available on industrial robots but can also be implemented on a custom-developed device. At this stage of the project, the Raspberry Pi 5 was employed as the communication interface in conjunction with the robot. A wired Ethernet-based link was employed to establish a physical connection between the two devices. Among the aforementioned requirements, the OPC UA data transmission solution was deemed to be the most appropriate.

A. OPC

The OPC (OLE for Process Control) is an industrial communication standard developed by the OPC Foundation. The inaugural relevant standard was published in 1996 [3]. The OPC Foundation was established by a number of prominent automation companies, including Siemens, Rockwell Automation, and ABB, with the objective of standardizing data exchange between industrial devices. The underlying technology, OLE (Object Linking and Embedding), was developed by Microsoft in the early 1990s [4].

The objective of OLE is to facilitate the sharing and integration of data and objects among disparate applications within the Windows environment. The OPC standard leverages the OLE (Object Linking and Embedding) and COM (Component Object Model)/DCOM (Distributed Component Object Model) technologies to facilitate the exchange of data between industrial automation devices and software [5], [6]. The employment of OLE technology by OPC facilitates the interoperability of devices produced by disparate manufacturers, thereby obviating the necessity for the development of custom drivers for each device.

B. OPC UA

OPC UA (OPC Unified Architecture) is a platform-independent, service-oriented architecture developed by the OPC Foundation for industrial automation. OPC UA represents an advancement over OPC Classic, addressing the shortcomings of its predecessor while providing a unified, secure, and reliable communication framework between industrial devices and software. The inaugural version of OPC UA was released in 2006 [7], [8].

Upon examination of the aforementioned requirements from the perspective of OPC UA, the following can be established:

- **Platform independence:** OPC UA is platform-independent and can run on various operating systems such as Windows, Linux, macOS, and embedded systems. This is a particularly important capability when using a Raspberry Pi.
- **Advanced security features:** OPC UA has advanced security features, including data encryption, signatures, authentication, and access control. Therefore, it meets high-level security requirements.
- **Support for complex data structures:** It supports complex data structures and object-oriented data modeling, which allows for richer and more semantic data exchange. Thus, complex data structures used in robotics also become accessible.
- **Support for various network protocols:** OPC UA supports various network protocols such as TCP/IP or HTTPS. Therefore, a wired TCP/IP connection provides an appropriate basis for communication.

II. System architecture

In the test system, a FANUC ER-4iA robot configured for educational purposes served as the industrial robot, and the middleware device running the test software was a Raspberry Pi 5. A wired TCP/IP connection was established between the two units. The robot's R-30iB Mate Plus controller possessed the necessary software to implement the OPC UA connection. This option is available on all currently marketed FANUC industrial robots, so it can be stated that the test environment modeled a state realized in industry. On the other side of the data transmission channel provided by the middleware device is the augmented reality device, represented by a Microsoft HoloLens 2. Figure 1 illustrates the system architecture of the communication between the robot and the middleware device.

The objective of the experiment was to demonstrate that a Python-based OPC UA client installed on a Raspberry Pi device can successfully read data from an industrial robot via an OPC UA connection.

In order to commence the test, the following prerequisites must be met:

- The OPC UA server on the industrial robot is enabled and correctly configured. A fixed IP address and port 4840 have been established.
- The robot and the Raspberry Pi are situated on the same network and are thus able to communicate with one another.
- The Raspberry Pi device has been equipped with both the Python 3 operating system and the OPC UA client library.

The initial phase of the test entailed mapping the data accessible within the FANUC robot and aligning the data formats. The objective was to extract data from the Position Register and other coordinate information. Furthermore, it was of interest to explore the full range of accessible data content. By initially reading the Position Register data, in addition to the position and orientation data stored in the robot's register, the configuration

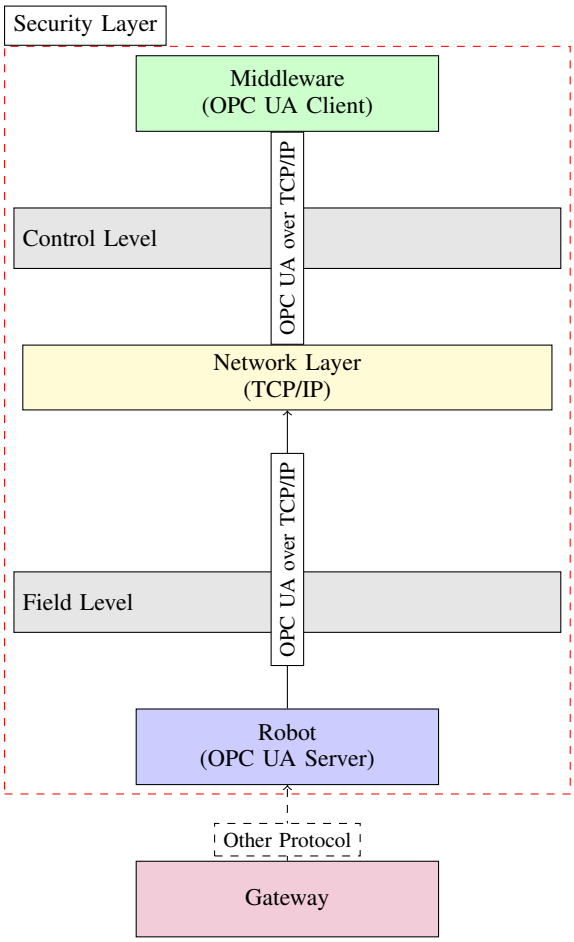


Fig. 1. Simplified OPC UA System Architecture with TCP/IP Network Layer

description, current tool and user coordinate system data, and the joint coordinates of all the robot's axes became accessible [**?**].

The data structure is shown in Figure 2.

III. The Result of the Test Process

The test results demonstrated the viability of the solution and provided the opportunity to develop the availability of all the data to be transferred to Hololens. The principal data yielded from the trial indicated that the solution was efficacious.

The data searched and read in the robot included the following:

- Position Registers
- User Frames
- Tool Frames
- DCS Area and Object data
- Payload geometry and inertia centre data

All of the aforementioned data is stored in the robot control system as system variables. It is important to note

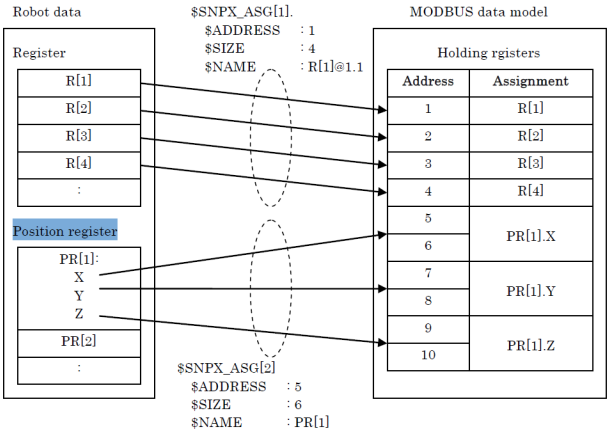


Fig. 2. Example of Holding Register assignment

that the read and write access to these system variables is not uniform. Moreover, the manufacturer restricts access to certain data for security reasons. Nevertheless, the issue of data access can be resolved through the utilisation of a range of supplementary programs and methodologies. To illustrate, data pertaining to position registers is readily available via the `$SNPX_ASG` system variable.

It is possible to ensure straightforward access to coordinate system data, such as that pertaining to user frames or tool frames, by copying the data to the position register using a program run on the robot and transmitting it via the OPC UA connection. This can be achieved through the utilisation of the PR[1] = UFRAME[1] instruction. The data pertaining to user frames is accessible via the `$MNUFRAME` system variable, whereas the data associated with tool frames can be obtained through the `$MNUTOOL` system variable. In contrast, the payload data is encapsulated within the `$PARAM_GROUP[group number].$payload` construct and its associated system variables, Fig. 3.

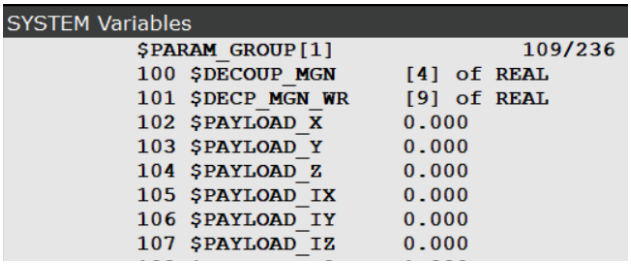


Fig. 3. System variables of the robot's payload

A. Operation of data processing

It is necessary to convert the read Position Register data to the appropriate format prior to its transmission to Hololens. The data received via OPC connection contains certain parameters of the Position Register, including X, Y, and Z, arranged in two consecutive byte sequences. The first 32 bits represent the integer part, while the second 32 bits represent the fractional part. These must be converted pairwise to floating point format to ensure suitability for subsequent mathematical operations and display. The Python program code for the underlying formula is shown in Figure 4.

```
if x1 < 0:
    eredmeny = x0 - (0.1 ** (len(str(abs(x1)))) * abs(x1))
else:
    eredmeny = x0 + (0.1 ** (len(str(abs(x1)))) * abs(x1))
```

Fig. 4. The equation for processing the data

Once all the requisite data has been converted, it can be stored and transferred to the Hololens device.

B. Storage and transmission of data

According to the previously described procedure, the processed data must be stored before being transmitted to the HoloLens. This storage is performed using a JSON format file [9], [10].

In comparison to alternative formats such as XML, YAML, and CSV, JSON was selected as the optimal choice for data storage and exchange. The verbosity of XML syntax would result in larger data files and increased processing requirements. Although XML offers a comprehensive framework for defining data, its considerable size renders it less suitable for use in environments where efficiency is of paramount importance and the capacities of middleware tools are limited. YAML was also considered, but due to the potential for parsing issues and the lack of standardisation, it was not deemed sufficiently reliable in an industrial environment where data consistency and reliability are of paramount importance. Although CSV is an efficient method of storing tabular data, it lacks the hierarchical data structure required to represent complex information, such as multidimensional position data for a robot.

The use of JSON was determined during the development of the system architecture. The format provides the ability to store, transmit, and validate data to ensure appropriate data security. JSON (JavaScript Object Notation) is a lightweight data interchange format developed in 2001 and standardized internationally in 2017 [11], [12]. The purpose of the format is to store and transmit data in a simple and easily readable manner while ensuring efficient machine processing.

A crucial element of the JSON data structure is the array, which stores data in an ordered list format. This form is particularly suitable for storing data read from the robot, as well as data conditioned by the middleware device. The JSON format is suitable for transmitting data over a network, offering an excellent solution for communication with the HoloLens.

To ensure the veracity of the data, the JSON Schema system is also intended for future deployment. This will facilitate the assessment of data integrity and correctness at both the point of origin and the point of reception [13].

Figure 5 illustrates the test phase of the system. The data displayed on the computer depicts the communication of the middleware device. The robot is engaged in a process of back-checking the accuracy of the read and subsequently conditioned data. Any manual modification of the data in the robot is promptly validated on the computer screen.

Fig. 5. Test environment at work

IV. Conclusion

The integration of industrial robots with augmented reality systems presents new opportunities for improving manufacturing efficiency and versatility. In this work, we demonstrated the use of a FANUC ER-4iA educational

robot with a Raspberry Pi middleware device to achieve efficient data exchange via OPC UA, and its subsequent utilization in a Microsoft HoloLens 2 system. The use of OPC UA for communication provides a secure, standardized, and platform-independent means of data exchange between the robot and the middleware, while JSON serves as an effective format for data storage, transmission, and validation.

The advantages of using JSON include its lightweight nature and human-readable format, which made it an optimal choice for storing and transmitting complex robot data. Moreover, the use of a Raspberry Pi middleware device showcased the flexibility of OPC UA, enabling connectivity between different systems with varying communication requirements.

The successful integration and data transmission demonstrate the feasibility of using a low-cost, versatile middleware solution to bridge industrial robotics and advanced visualization systems like augmented reality. The results show that the system architecture presented is applicable in industrial settings, potentially enabling enhanced user interaction and process visualization through augmented reality.

Future work could involve further optimization of the communication layer, investigating the potential of real-time data processing, and exploring other augmented reality applications for robotic systems to further expand the capabilities of such integrated solutions.

REFERENCES

[1] OASIS, *MQTT Version 3.1.1*, OASIS, 2014. [Online]. Available: https://docs.oasis-open.org/mqtt/mqtt/v3.1.1/os/mqtt-v3.1.1-os.html

[2] A. P.S., S. Dilip Kumar, and V. K.R., "MQTT implementations, open issues, and challenges: A detailed comparisonand survey," *International Journal of Sensors, Wireless Communications and Control*, vol. 12, no. 8, pp. 553–576.

[3] OPC Foundation, *OPC Data Access Specification Version 1.0*, OPC Foundation, 1996.

[4] K. Brockschmidt, *Inside OLE*, second edition ed. Microsoft Press, 1995.

[5] D. Box, *Essential COM*. Addison-Wesley Professional, 1998.

[6] G. Eddon and H. Eddon, *Inside Distributed COM*. Microsoft Press, 1998.

[7] OPC Foundation, *OPC Unified Architecture Specification*, OPC Foundation, 2006. [Online]. Available: https://opcfoundation.org

[8] A. Busboom, "Automated generation of OPC UA information models — a review and outlook," *Journal of Industrial Information Integration*, vol. 39, p. 13.

[9] D. Crockford, "Json: A lightweight data interchange format," 2001, accessed: YYYY-MM-DD. [Online]. Available: https://www.json.org/json-en.html

[10] T. Bray, "The javascript object notation (json) data interchange format," Internet Engineering Task Force, Request for Comments 8259, Dec. 2017. [Online]. Available: https://tools.ietf.org/html/rfc8259

[11] *The JSON Data Interchange Format*, ECMA International Std. ECMA-404, 2013. [Online]. Available: https://www.ecma-international.org/publications-and-standards/standards/ecma-404/

[12] *Information technology – The JSON data interchange format*, ISO/IEC Std. ISO/IEC 21 778:2017, 2017. [Online]. Available: https://www.iso.org/standard/71616.html

[13] F. Pezoa, J. L. Reutter, F. Suarez, M. Ugarte, and D. Vrgoč, "Foundations of JSON schema," in *Proceedings of the 25th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, pp. 263–273.

Dynamic Damper Control Extension for Two-wheeled Applications

Péter Gyenis
Obuda University
Alba Regia Technical Faculty
Székesfehérvár, Hungary
gyenis.peter@stud.uni-obuda.hu

Bertalan Beszédes
Obuda University
Alba Regia Technical Faculty
Székesfehérvár, Hungary
beszedes.bertalan@uni-obuda.hu
https://orcid.org/0000-0002-9350-1802

***Abstract*— This paper presents the development of a dynamic control system that expends the functionality of the shock absorber of a two-wheeled, human propulsion vehicle, which adapts to road conditions, both in urban and off-road environments. The road surface and tilt dependent dynamic damper control system reduces vibrations and enhances stability, resulting in a smoother and safer ride, with increases efficiency and versatility.**

***Keywords*— vehiche dinamics, shock absorber, dinamic damping, bicicle control, bicicle dinamics, bicycle frame, anti-squat, dynamically stabilized vehicle, personal mobility vehicle**

I. INTRODUCTION

This article aims to improve the suspension efficiency of mountain bikes, combining the power delivered by turning the pedal to the ground as effectively as possible, as well as optimizing ground tracking while rolling and absorbing large shocks. The article's documentation first details the elements of the suspension, as well as examines the kinematic roots of the problem. It details the possible options for improving efficiency, and then examines the solutions currently available on the market. Finally, it discusses the implementation, with the aspects of the selection of the components used, the description of the chosen units, and the structure and operation of the control program.

A. Elements of the suspension

The spring and air spring are undamped energy reservoirs that absorb/store/resolve the vertical forces acting on the suspension. It is important that, in order to keep the correct ground tracking and dynamic ride height in the right range, when dimensioning, it is worth choosing the spring stiffness so that the suspension has 25-30% seat in a static state when loaded with the cyclist and his equipment.

The two-chamber oil-filled damper, which is used to limit oscillation on the one hand, and to increase winding efficiency by partially or completely closing it on the other. In most cases, it has two (compression and rebound) or four (fast/slow compression and rebound) adjustment valves. In this work, slow compression is controlled.

The materials used in bicycle frame production each come with specific advantages and disadvantages. Steel offers high flexibility, driving comfort, ease of repair, and a low price. However, it also has significant drawbacks, including high weight, poor corrosion resistance, and difficulties when used with full-suspension frames. Aluminum alloy is favored for its low weight, low price, machinability, malleability, and corrosion resistance. Despite these benefits, it compromises on driving comfort, is difficult to repair, and has stiffness issues.

Magnesium alloy is advantageous due to its low weight and ability to absorb vibrations well. On the downside, it has poor corrosion resistance, a low lifetime, and a high price. Titanium alloy stands out for its resistance, comfort, durability, stiffness, and the fact that it does not require surface treatment. However, it is expensive, difficult to manufacture, challenging to weld, and heavier compared to aluminum.

Carbon fiber composite excels with its very low weight, stiffness, capability of forming complicated shapes during production, repairability, and vibration absorption. Nevertheless, it comes with a high price, difficulty in processing and welding, and a higher weight compared to aluminum.

In addition to these common materials, bamboo and cardboard have also been explored for bicycle frames. Bamboo offers sustainability, shock absorption, and a unique aesthetic, but it can be inconsistent in quality and requires careful treatment to prevent decay. Cardboard, while innovative and eco-friendly, is less durable and poses challenges in terms of weather resistance and longevity.

The swing arms are used to connect the rigid frame of the bicycle and the wheel, namely by allowing the wheel to move in its vertical degree of freedom. They can be found in several designs, there are solutions with one pivot point, multiple pivot points and even sliding shafts. The project deals with a bicycle with a five-pivot four-link "horst link" swing arm system [1], see Fig. 1. In terms of rolling efficiency, its most important kinematic feature is anti-squat [2]. Table 1. summarizes the advantages, disadvantages and manufacturers of different types of rocker arms [3].

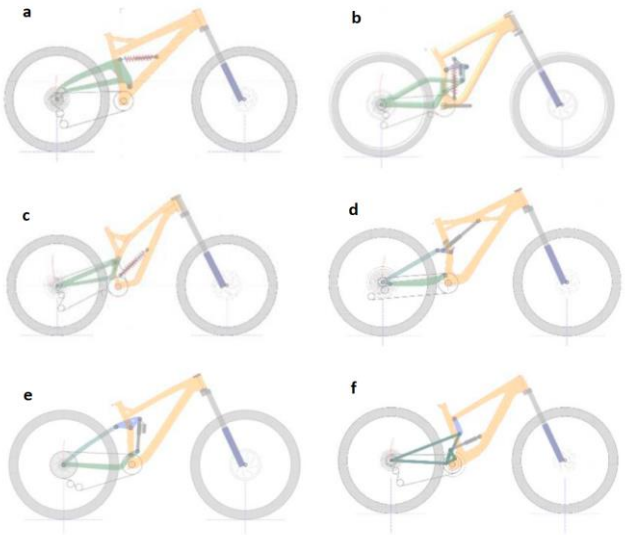


Figure 1: Single pivot (a); Single pivot with additional leverage (b); Single pivot with highly placed pin (c); Four pivot system – Horst Link (d); Four pivot system – Split Pivot (e); Four pivot system – VPP (f)

Suspension System	Advantages	Disadvantages	Company
One pivot	very simple, minimum service requirements	bad pedalling response moving is affecting by braking	Dewilwork, Morewood, Santa Cruz, Canondale

One pivot with levering	better regulation of the applied force on the damper	moving is affecting by braking, moving of rear tringle when pedalling	Merida, Kona, Trek, Transition, Scott
One-pivot High Pivot	eliminated impact of braking to suspension	long length of chain, higher weight, used just with downhill bikes	Commencal, Scott
Four-pivot Horst Link	simple system, almost no pedalling squat	suspension is sensitive to small terrain unevenness, affected by braking	Canyon, Radon, Specialized, Norco
Four-pivot Split pivot	the best drivability, a small drop of the rear triangle when pedalling, removes brake squat	higher weight, patent litigation	Pivot, Morewood, Trek
Four-pivot Twin link	balanced movement, suspension flexibility, almost no suspension moving when pedalling	higher weight, higher maintenance requirements	Santa Cruz, Ibis, Giant

Table 1: Comparison of suspension systems

II. THEORETICAL BACKGROUND

A. *Factors affecting winding efficiency*

The power delivered by the cyclist when pedaling is not a steady circular motion, but a stomping motion. In this way, instead of generating pure torque, some of the power is converted into heat by the compression of the suspension by the damping. The effectiveness is further worsened by the fact that, in the absence of a snap pedal, in the second half of the pedal stroke, some resistance must also be performed with the foot in order not to lose sole-pedal contact.

Anti-squat is a ratio that determines whether the force transmitted to the rear sprocket by the chain during pedaling compresses or stretches the suspension. To calculate it, you must first project a straight line between the intersection of the chain line/swing arm line and the ground contact point of the rear wheel, then take the vertical distance of this line from the ground in the axle of the front wheel, and then divide it by the vertical distance of the center of mass to the ground. Since this value varies from the chain line, the suspension seat and the center of mass, it is impossible to tune in such a way as to compensate for the vertical displacement generated when pedaling.

Since pedaling efficiency has a physiological limit, and the use of anti-squat cannot be optimal due to the variables, a significant increase in efficiency can be achieved by closing the valves of the shock absorber. Since the average cyclist turns the crank at 60-90 rpm, this falls into the low stem speed range. Limitation of this movement can be achieved either by completely closing the damping or by using a highly digressive ($\alpha<1$) damping curve, see Fig. 2. In order to absorb larger shocks expected during free fall or downhill, in these cases we need a linear ($\alpha=1$) or slightly progressive ($\alpha>1$) characteristic. Where α is the damping force – velocity relation for various values of exponent of velocity. The solution is the real-time control of the damping characteristic according to the conditions [5-10].

B. *Commercially available solutions*

Currently, the two largest brands of suspension struts and forks offer active bicycle suspension solutions. Both manufacturers launched their products at the end of 2019, which were initially only available as a solution built into the bike by the manufacturer. The starting price of these bicycles is between 5100 and 7700 euros, which compared to products with traditional non-controlled suspension but with the same equipment, the premium for active control means a difference of roughly 2000 euros. At the end of 2022, RockShox made its product available as an aftermarket option, which includes an active spring member, an active telescope and a pedal rotation sensor, with a suggested retail price of 3,130 euros.

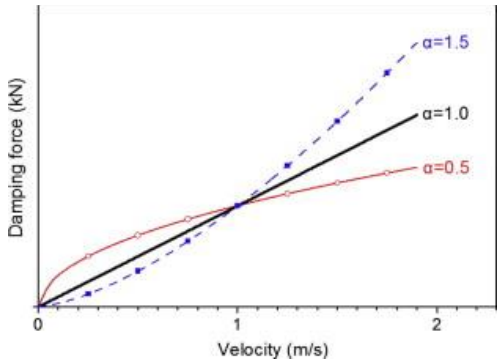


Figure 2: Flowchart of the alarm system [4]

The units are connected to each other via wireless communication. Due to the closed system, detailed information about its exact operation cannot be found. Based on the available information, the system coordinates both spring members and operates with the valves open in the normal state. When rolling on smooth terrain (service roads, asphalt roads), it closes the system completely for maximum efficiency, and provides partial closure on uneven terrain. When going downhill and free-falling, it fully opens the damping valves.

Fox's solution is based on the acceleration measurement of the central unit and the front/rear unsprung elements, which are connected by wire. Exact information is not available here either, but Live Valve works with closed rear and open front valves by default. The opening of the rear member is controlled by the central unit based on the shocks detected by the front wheel, and the system opens completely in free fall. The first telescope is closed only when the central unit detects from the rhythmic movement that the cyclist is riding with strong pedaling movements, standing up from the saddle.

III. REALIZATION OF A UNIQUE IMPLEMENTATION

A. *Mechatronical design*

For the implementation, a uniquely designed bicycle frame calculated with a static center of mass was designed, shown in Figure 3. The Anti-squat graph displayed as a function of the relevant wheel path is shown in Figure 4.

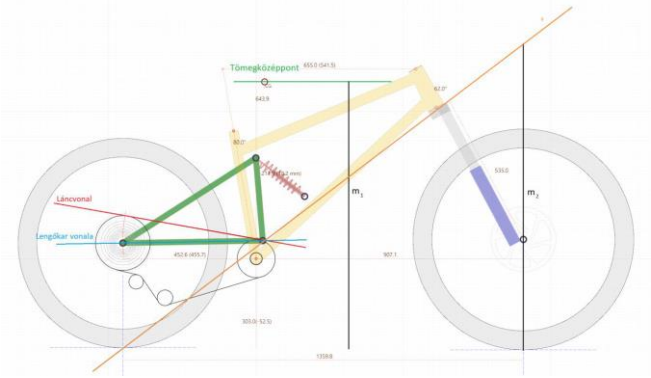


Figure 3: Bicycle frame to calculate Anti-squat

To implement the prototype, the central control unit is an Arm Cortex-M4F microcontroller-based development board (Arduino Nano 33 BLE), which has the possibility of implementing a bluetooth wireless connection and a matched

antenna, as well as an integrated 9-axis IMU (Inertial measurement Unit) sensor (LSM9DS1). The digital acceleration sensor, gyroscope and magnetic field sensor IC manufactured by ST Microelectronics (LSM9DS1:) was matched to this. The latter has four measurement ranges, the smallest of which (±2g) is sufficient for current use.

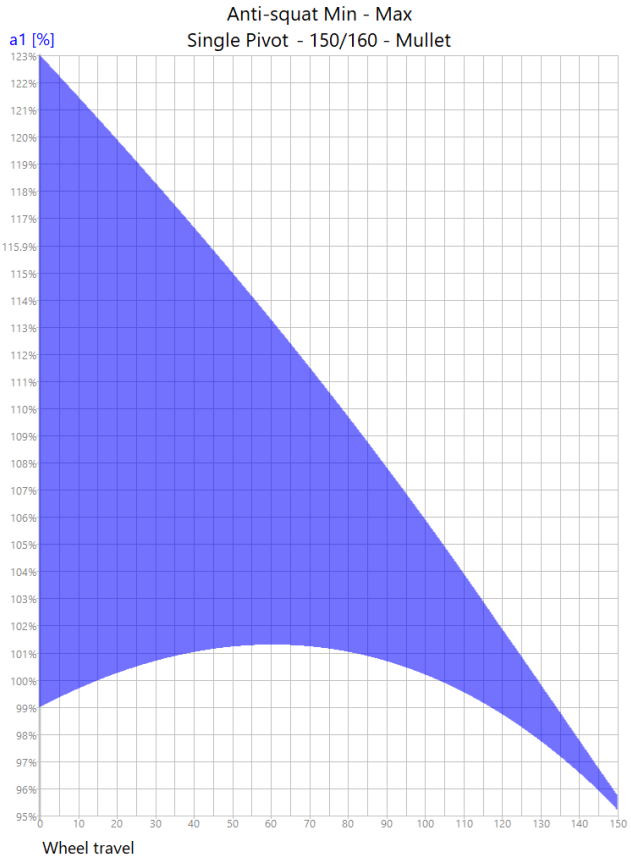


Figure 4: Anti-squat graph calculated with static center of mass

A 28BYJ-48 geared stepper motor was selected as an intervention during the mechanical construction of the prototype, the advantages of which are the small size and weight well suited to the application, mechanical self-support, energy efficiency, low supply voltage, availability and low initial cost. The connection to the microcontroller is provided by a power driver IC type ULN2003. Other additional components include a microswitch for recording end positions, current limiting resistors, pull-up resistors and connectors.

B. Electronics design

A development board shield printed circuit was designed for the prototype, which includes the necessary wiring, the USB power input, the stepper motor controller and limit switches. For the PCB design, the goal was to create an elongated shape for easier placement on the bicycle frame, see Figure 5. Similar hardware close, microcontroller based firmware solutions can be seen in [11-16]. The microcontroller-based solutions presented here can also be applied well in technical frontier areas [17-22].

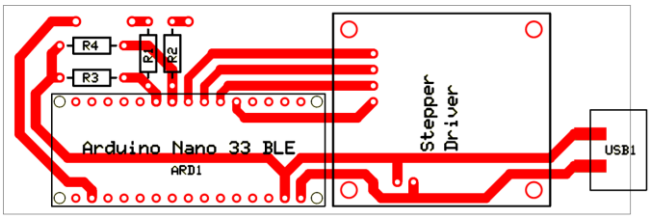


Figure 5: PCB design

C. Mechanical attachment

The attachment of the 28BYDJ-48 stepper motor to the spring strut (DT Swiss R 414) was solved by inserting a 3D printed joint made of ABS plastic. This intermediate piece can be attached to the recess separating the positive and negative air chambers of the spring strut with the help of a two-membered ring, and its shape that rests on the positive chamber helps the stable fixation.

The flattened shaft of the stepper motor and the hexagonal shaft of the strut were also connected with a 3D printed element made of ABS plastic (see Fig. 6.), since the low torque does not require the use of a more resistant material.

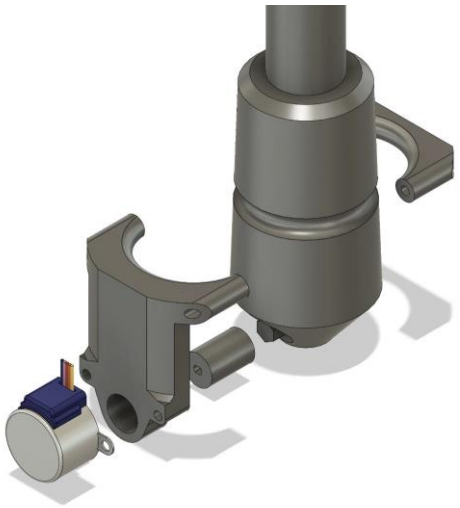


Figure 6: Intervenor isometric exploded view

D. Software implementation

In the program, after calling the IMU and stepper motor libraries, first initialize the stepper motor named m1 "AccelStepper m1(4, 5, 3, 4, 2);", where the first variable is the number of controlled PINs, the 2- 4 variables are the addressing of the PINs. In the next step, the software sets the inputs of the two limit switches as inputs. Finally, it sets the maximum allowable stepping speed, current speed and acceleration of the stepper motor. The values used were determined empirically. The main program first enables the IMU and the stepper motor m1, then creates the variables used: zerog - reference vertical, zerom1 - reference motor position, m1trg - stepper motor target position, min, max - minimum and maximum end positions.

The next step is the calibration of the end positions, which the program performs by writing the return value of the "calibrate" subroutine into the variable. After that, it calculates the middle position of the stepper motor by the quotient of the two end positions. It then picks up the current vertical position by writing the return value of the "m1target" subroutine to the "zero" variable.

After it have finished setting the basic conditions, the program enters a loop where it determines the desired motor position and then sends the motor there with the m1go subroutine, follow on Fig. 7. Hysteresis is programmed into the m1go subroutine, which keeps the control electronics of the stepper motor asleep by turning off the drive stage in case of small changes.

Communication between all elements of the system must be reliable and secure. It is proposed to regularly checking the system to ensure all elements function correctly, also setting up automatic notifications and alerts for errors, intrusion attempts, or other suspicious activities [23-26].

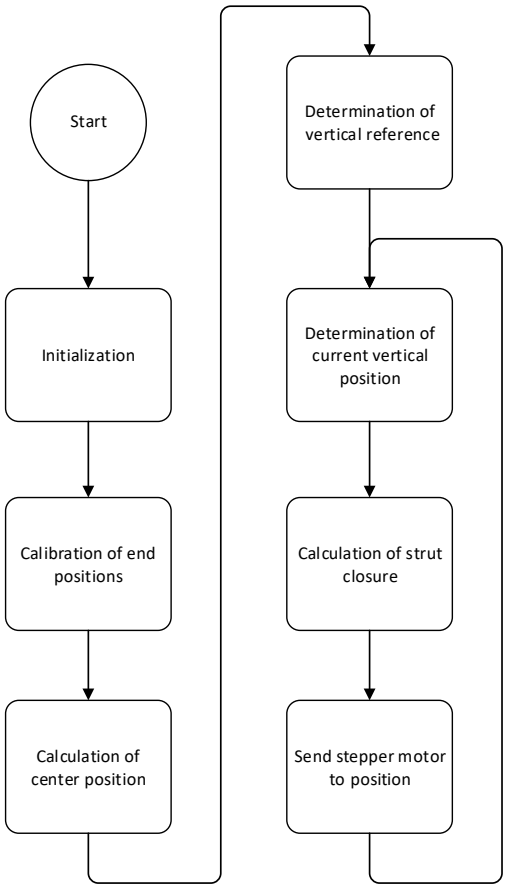


Figure 8: Flow chart describing the operation

CONCLUSION

The integration of dynamic damper control with human propulsion in a two-wheeled vehicle significantly enhances stability, reduces vibrations, and improves rider comfort. The development of a 3D model and functional prototype, along with a custom microcontroller program, validated the theoretical benefits and preliminary expectations. This innovative approach represents a substantial advancement in personal transportation, offering more efficient, versatile, and comfortable travel solutions for urban commuting and recreational use.

ACKNOWLEDGMENT

The author would like to thank all the Obuda University Alba Regia Technical Faculty staff and member that provide help and assistance throughout the project completion.

REFERENCES

[1] Michal Bartoš, Martin Bohušík, Vladimír Tlach, Tomáš Dodok, Ivan Zajačko. Increasing the strength of full suspension bicycle frame by component modification within maintaining a bike geometry. Transportation Research Procedia. Volume 74. 2023. pp. 555-562. ISSN 2352-1465. https://doi.org/10.1016/j.trpro.2023.11.181.

[2] Nielens, H., Lejeune, T. Bicycle Shock Absorption Systems and Energy Expended by the Cyclist. *Sports Med* 34, 71–80 (2004). https://doi.org/10.2165/00007256-200434020-00001

[3] Bulej, Vladimír, Ivan Kuric, Milan Sága, Milan Vaško, Zuzana Ságová, Michal Bartoš, and Stanislaw Legutko. 2022. "Analysis of Symmetrical/Asymmetrical Loading Influence of the Full-Suspension Downhill Bicycle's Frame on the Crack Failure Formation at a Critical Point during Different Driving Scenarios and Design Improvement" *Symmetry* 14, no. 2: 255. https://doi.org/10.3390/sym14020255

[4] George D. Hatzigeorgiou, Nikos G. Pnevmatikos. Maximum damping forces for structures with viscous dampers under near-source earthquakes. Engineering. Structures. Volume 68. 2014. Pages 1-13. ISSN 0141-0296. https://doi.org/10.1016/j.engstruct.2014.02.036.

[5] P. De Filippi, M. Corno, M. Tanelli and S. M. Savaresi, "Single-Sensor Control Strategies for Semi-Active Steering Damper Control in Two-Wheeled Vehicles," in IEEE Transactions on Vehicular Technology, vol. 61, no. 2, pp. 813-820, Feb. 2012, doi: 10.1109/TVT.2011.2180031.

[6] P. De Filippi, M. Tanelli, M. Corno, S. M. Savaresi and L. Fabbri, "Semi-Active Steering Damper Control in Two-Wheeled Vehicles," in IEEE Transactions on Control Systems Technology, vol. 19, no. 5, pp. 1003-1020, Sept. 2011, doi: 10.1109/TCST.2010.2070068.

[7] Devikiran P, Puneet N, Hegale A, Kumar H. Design and development of MR damper for two wheeler application and Kwok model parameters tuning for designed damper. Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering. 2022;236(7):1595-1606. doi:10.1177/09544070211036317

[8] Lamri Nehaoua, Hichem Arioui, Nicolas Seguy, Saïd Mammar. Dynamic modelling of a two-wheeled vehicle: Jourdain formalism. Vehicle System Dynamics, 2013, 51 (5), pp.648- 670. 10.1080/00423114.2012.762536. hal-00786265

[9] Mara Tanelli, Matteo Corno, Ivo Boniolo, Sergio M. Savaresi. Active braking control of two-wheeled vehicles on curves. International Journal of Vehicle Autonomous SystemsVol. 7, No. 3-4. 2010. pp 243-269. https://doi.org/10.1504/IJVAS.2009.033263 ISSN: 1471-0226

[10] Chihiro Nakagawa, Kimihiko Nakano, Yoshihiro Suda, Ryuzo Hayashi. Stability of the Dynamically Stabilized Two-Wheeled Vehicle Traveling on a Rough Road. ournal of Mechanical Systems for Transportation and Logistics. Vol. 2, No. 1, 2009. DOI: 10.1299/jmtl.2.78

[11] Mosavi Amirhosein, Bertalan Beszedes, Imre Felde, Nadai Laszlo, Gorji Nima E. Electrical characterization of CIGS thin-film solar cells by two- and four-wire probe technique. MODERN PHYSICS LETTERS B 2020 p. 2050102 , 16 p. (2020)

[12] György Györök, Baklanov Alexander E, Bertalan Beszédes. Extension of Nodal Voltage Method with the Thermosensing. In: Orosz, Gábor Tamás (szerk.) AIS 2017 - 12th International Symposium on Applied Informatics and Related Areas organized in the frame of Hungarian Science Festival 2017 by Óbuda University : Proceedings. Székesfehérvár, Magyarország : Óbudai Egyetem (2017) 204 p. pp. 201-204. , 4 p.

[13] György Györök, András Dávid, Nikoletta Tolner, Bertalan Beszédes, Dániel Cseh. Supervision of the operation of digital circuits by Embedded Microcontroller. In: Orosz, Gábor (szerk.) AIS 2018 - 13th International Symposium on Applied Informatics and Related Areas. Székesfehérvár, Magyarország : Óbudai Egyetem, Alba Regia Műszaki Kar (2018) pp. 29-35. , 7 p.

[14] Györök György, Tihomir Trifonov, Alexander E. Baklanov, Bertalan Beszédes, Svetlana V. Grigoryeva, Aizhan Zhaparova. A Special Robust Solution for Battery Based Power Supply. In: Orosz, Gábor Tamás (szerk.) 11th International Symposium on Applied Informatics and Related Areas (AIS 2016). Székesfehérvár, Magyarország : Óbudai Egyetem (2016) pp. 32-35. , 4 p.

[15] Attila Sáfár, Bertalan Beszédes. Educational Aspects of a Modular Power Management System. In: Orosz, Gábor Tamás (szerk.) AIS 2019 : 14th International Symposium on Applied Informatics and Related Areas organized in the frame of Hungarian Science Festival 2019 by Óbuda University. Székesfehérvár, Magyarország : Óbudai Egyetem (2019) pp. 163-166. , 4 p.

[16] György Györök, Bertalan Beszédes. Artificial Education Process Environment for Embedded Systems. In: Orosz, Gábor Tamás (szerk.) 9th International Symposium on Applied Informatics and Related Areas - AIS2014. Székesfehérvár, Magyarország : Óbudai Egyetem (2014) pp. 37-42. , 6 p.

[17] I. Bendiák and T. Sándor, "Possible Ways of Measuring and Calculating Waste Heat from a Machine Diagnostic Approach," 2023 IEEE 6th International Conference and Workshop Óbuda on Electrical and Power Engineering (CANDO-EPE), Budapest, Hungary, 2023, pp. 000259-000266, doi: 10.1109/CANDO-EPE60507.2023.10417972.

[18] I. Bendiák and S. Semperger, "Simplified Predictive Strategy of Mechanical Life Cycle Model in Three-Phase Asynchronous Motor," 2023 IEEE 6th International Conference and Workshop Óbuda on Electrical and Power Engineering (CANDO-EPE), Budapest, Hungary, 2023, pp. 000267-000274, doi: 10.1109/CANDO-EPE60507.2023.10417977.

[19] Alexander Baklanov, Svetlana Grigoryeva, György Györök. Control of LED Lighting Equipment with Robustness Elements, Acta Polytechnica Hungarica (1785-8860 1785-8860): 15 3 pp 105-119 (2016)

[20] G. Gyorok and M. Mako, "Configuration of EEG input-unit by electric circuit evolution," 2005 IEEE International Conference on Intelligent Engineering Systems, 2005. INES '05., Spain, 2005, pp. 255-258, doi: 10.1109/INES.2005.1555168.

[21] Aizhan Zhaparova, Dimitry Titov, Alexander Y. Baklanov, Gyorgy Gyorok. Study of the Effectiveness of Switching-on LED Illumination Devices and the Use of Low Voltage System in Lighting. Acta Polytechnica Hungarica (1785-8860 1785-8860): 12 5 pp 71-80 (2015)

[22] S. Kimathi, B. Lantos. Modelling and Attitude Control of an Agile Fixed Wing UAV based on Nonlinear Dynamic Inversion, Periodica Polytechnica Electrical Engineering and Computer Science, Vol 66(3), 2022, pp. 227-235

[23] S. Kimathi, B. Lantos. Quaternion based attitude control of a maneuvering fixed wing UAV, IEEE 26th International Conference on Intelligent Engineering Systems (INES), pp. 261-266, 2022.

[24] Thermo-dynamic Cycle Computation of a Micro Turbojet Engine. Fozo, L; Andoga, R and Kovacs, R. 17th IEEE International Symposium on Computational Intelligence and Informatics (CINTI). 2016 17TH IEEE International Symposium on Computational Intelligence and Informatics (CINTI 2016) , pp.75-79

[25] Z. Varga and E. Racz, 'Machine Learning Analysis on the Performance of Dye-Sensitized Solar Cell—Thermoelectric Generator Hybrid System', Energies, vol. 15, no. 19, p. 7222, Oct. 2022, doi: 10.3390/en15197222.

[26] Z. Varga and E. Racz, 'A Numerical Model and a Code Development for Photogeneration Rate Calculation for a Dye Sensitized Solar Cell', in 2022 IEEE 16th International Symposium on Applied Computational Intelligence and Informatics (SACI), Timisoara, Romania: IEEE, May 2022, pp. 000341–000346. doi: 10.1109/SACI55618.2022.9919499.

Software Simulation of a Microcontroller-Based Wheeled Mobile Robot

Alexander Krasavin
School of Digital Technologies and Artificial Intelligence
D Serikbayev East Kazakhstan Technical University
Ust-Kamenogorsk, Kazakhstan
akrassavin@ektu.kz

Darya Alontseva
School of Digital Technologies and Artificial Intelligence
D Serikbayev East Kazakhstan Technical University
Ust-Kamenogorsk, Kazakhstan
dalontseva@ektu.kz

Gaukhar Nazenova
School of Digital Technologies and Artificial Intelligence
D Serikbayev East Kazakhstan Technical University
Ust-Kamenogorsk, Kazakhstan
nazenova.g@edu.ektu.kz

***Abstract* — Currently, mobile wheeled robots are widely used as platforms for education and research in automatic control, robotics, and related fields. Despite the simplicity of kinematic models, dynamic models of mobile robots are significantly nonlinear. This characteristic, combined with the simplicity of mechanical design, makes them a convenient platform for experimentation in the field of nonlinear automated control. Prior to physical experiments, computer simulations of the robot's control system are usually conducted. MATLAB/Simulink is most commonly used to create software simulators for wheeled mobile robots (WMR). These simulators model continuous systems, with control laws described by analytically defined expressions. However, in reality, a mobile robot's control system is typically a microcontroller (MCU) or a single-board computer, and the software-implemented system controller works with discrete sensor signals and generates discrete control signals. This paper describes a novel framework for a wheeled mobile robot software simulation in which the robot's controller is modeled as a virtual microcontroller. Control algorithms implemented as code of some simulator Application Programming Interface (API) functions, which can be considered as virtual MCU interrupt handlers.**

Keywords — wheeled mobile robot modeling, wheeled mobile robot software simulator, kinematic model, dynamic model, virtual MCU.

I. Introduction

Nowadays, mobile robots play a key role in the development of robotics, finding applications in a wide range of fields - from industry to artificial intelligence research. Among the many types of mobile robots, two-wheeled mobile robots stand out due to their simple design and broad range of capabilities. This makes them popular subjects for the study and development of new control methods [1].

The motivation for choosing a two-wheeled robot in this article is based on several factors. First, the market offers a wide range of affordable chassis for such robots, making them easy to use in educational and research projects [2]. Second, despite their apparent simplicity, controlling a WMR is a complex task due to the nonlinearity of its kinematic model [3]. This combination of affordability and complex dynamics makes two-wheeled mobile robots a popular platform for research, particularly in the development of new control systems, such as inertial navigation systems using machine learning [4]. To ensure the reliability and fault tolerance of mobile robot control systems, it is important to consider redundancy approaches such as duplicated control units and fault-tolerant power supplies. Thus, G. Györök and B. Beszedes in paper [5] describe methods for creating fault-tolerant power supply systems that can maintain continuous operation even when individual components fail. Additionally, the duplication of control units, as discussed by the authors in article [6], enables the implementation of fault-masking systems, thereby ensuring the stability and reliability of robot operation in dynamic and unpredictable environments.

One of the key challenges in developing control systems for mobile robots is the need to simulate their operation. Computer modeling allows for experiments and testing of algorithms without the need for physical prototyping, significantly saving time and resources [7]. Currently, there are many tools available for robot simulation, such as Simulink MATLAB, which are widely used for modeling manipulators and control systems. For example, V. Mudeng et al. use Simulink MATLAB to simulate the control system of a self-balancing robot with a Proportional–Integral–Derivative (PID) controller. The simulations involve mathematical modeling, deriving transfer functions, and simulating the robot's behavior based on the inverted pendulum model [8].

H. Xue et al. also use Simulink MATLAB to simulate their proposed adaptive control algorithm based on an improved grey wolf algorithm. It allows for testing robot control algorithms while accounting for temporal and external disturbances. In this context, Simulink serves as a platform for verifying algorithms, including the use of neural networks and intelligent control systems to adjust the robot's parameters in real-time [9].

However, analyzing the articles [8-10], it can be observed that the simulators used, despite their advantages, have significant drawbacks. First, there are high computational costs: Simulating complex systems, such as mobile robots considering all dynamic models and adaptive control algorithms, can require substantial computational resources, making simulations lengthy and energy-intensive. Second, there are real-time limitations: Simulink may not always accurately replicate the operation of a robot in real-time, especially if the system is highly nonlinear or involves delays in external disturbances. Third, there are challenges in integrating with real hardware: while Simulink provides a convenient platform for simulation, translating models to physical equipment can encounter compatibility issues and requires additional tuning, particularly when controlling real mobile robots.

The simulation of microcontroller-based control systems, such as those employed in wheeled mobile robots, highlights the importance of robust fault-tolerant software solutions,

which ensure system reliability even in the presence of errors or disturbances. The principles discussed in [11] can be directly applied to enhance the resilience of MCU-controlled robots, improving their performance in dynamic and unpredictable environments.

Thus, the simulators used in research articles allow for testing control laws described by continuous functions, but their capabilities are limited when it comes to implementation on microcontrollers, where data and control algorithms are often discrete. In this regard, our article proposes the simulation of a two-wheeled mobile robot controlled by a microcontroller in Python. Using Python, various levels of simulation can be created for the two-wheeled robot, ranging from simple 2D graphics to complex physical simulations using libraries such as PyBullet, ROS, or Gazebo. The choice of framework depends on the complexity of the task, the accuracy requirements of the simulation, and the available computational resources [12].

One of the goals of this work is to develop a software simulator for a WMR using a specialized framework that will allow for the simulation of microcontrollers and control systems. The simulation will include data processing from sensors and the execution of control algorithms, which are traditionally implemented through timer interrupt handling in microcontrollers. The proposed framework will provide an application programming interface for simulating data acquisition from sensors, managing time parameters, frequency, and cycle duration, thereby enabling a more accurate modeling of real systems.

This paper presents a prototype software simulator for the microcontroller control system of a Differential Drive Wheeled Mobile Robot (DDWMR), featuring several desirable attributes. The key innovation in our proposed simulation framework is the concept of a virtual microcontroller that interacts with the hardware environment simulator. The virtual microcontroller is implemented as a set of user-defined functions that simulate the interrupt handlers of the real microcontroller managing the DDWMR. During simulation, the hardware environment simulator invokes these functions, emulating hardware interrupts from timers that regulate the control system's timing, as well as interrupts generated by the microcontroller's capture blocks. The simulation system provides an API for the user to read data from virtual sensors generated by the hardware environment simulator while the virtual interrupt handler code runs, as well as to set control codes simulating robot motor control. The set of virtual interrupt handlers aligns with the event-driven microcontroller control application framework commonly used in embedded systems development.

II. Methods

A. Kinematic model and constraint equations

Examine the dynamics of a two-wheeled mobile robot that is capable of moving forward and rotating around its geometric center, as illustrated in Figure 1.

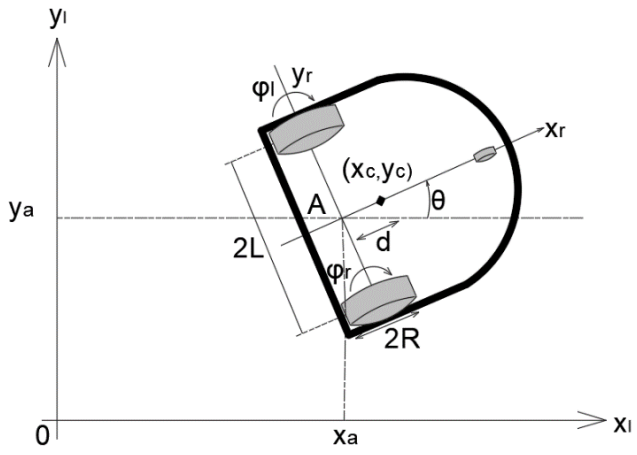


Fig. 1. A diagram of a Differential Drive Wheeled Mobile Robot

Let $\{X_I, Y_I\}$ represent a fixed Cartesian inertial coordinate system, and let $\{X_r, Y_r\}$ п denote the moving Cartesian coordinate system of the robot. The origin of the robot's coordinate system is located at point A, situated in the middle of the axis between the wheels. The center of mass of the robot is assumed to be positioned along the axis of symmetry, at a distance d from point A.

The motion of a differential-drive mobile robot is characterized by two nonholonomic constraint equations, which are derived from two main assumptions:

No lateral slipping: the robot can only move along a curve (forward and backward) and not sideways. In the robot's coordinate system, this condition implies that the velocity of the central point A is zero along the lateral axis, as indicated in expression (1):

$$\dot{y}_\alpha^r = 0 \tag{1}$$

Pure rolling constraint: the pure rolling constraint refers to the fact that each wheel maintains a single point of contact P with the ground. There is no slipping of the wheel along its longitudinal axis, and no skidding along its orthogonal axis. The velocities of the contact points in the robot's frame are related to the wheel speeds as follows (2):

$$\begin{cases} v_{pR} = R\dot{\varphi}_R \\ v_{pL} = R\dot{\varphi}_L \end{cases} \tag{2}$$

Thus, the three constraint equations can be expressed in the following matrix form:

$$\Lambda(q)\dot{q} = 0, \tag{3}$$

where

$$\Lambda(q) = \begin{bmatrix} -\sin\theta & \cos\theta & 0 & 0 & 0 \\ \cos\theta & \sin\theta & L & -R & 0 \\ \cos\theta & \sin\theta & -L & 0 & -R \end{bmatrix} \tag{4}$$

and

$$\dot{q} = \left[\dot{x}_\alpha\ \dot{y}_\alpha\ \dot{\theta}\ \dot{\varphi}_R\ \dot{\varphi}_L\right]^T \tag{5}$$

The kinematic model of a differential-drive mobile robot is characterized by its extreme simplicity and is described by the well-known equations (6) and (7) for the linear velocity of the robot's center of mass v and the angular velocity $\omega = \dot{\theta}$ of the robot.

$$v = \frac{v_R + v_L}{2} = R\frac{(\varphi_R + \varphi_L)}{2} \tag{6}$$

$$\omega = \frac{v_R - v_L}{2L} = \frac{(\dot{\varphi}_R - \dot{\varphi}_L)}{2} \quad (7)$$

The Lagrange-Euler equations for the DDWMR take the form (8):

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) + \frac{\partial L}{\partial q_i} = F - \Lambda^T(\boldsymbol{q})^{\lambda}, \quad (8)$$

where L=T-V is the Lagrangian function, T is the kinetic energy of the system, V is the potential energy of the system, qi are the generalized coordinates, F is the generalized force vector, Λ is the constraints matrix, and λ is the vector of Lagrange multipliers associated with the constraints. As demonstrated in reference [13], the equations in Lagrange form (8) are equivalent to the dynamics equation of the mobile robot (9):

$$\begin{cases} \left(m + \frac{2I_w}{R^2}\right)\dot{v} - m_c d\omega^2 = \frac{1}{R}(\tau_r + \tau_L) \\ \left(I + \frac{2L^2}{R^2}I_w\right)\dot{\omega} + m_c d\omega v = \frac{L}{R}(\tau_r - \tau_L), \end{cases} \quad (9)$$

where τ_R and τ_L are the moments of frictional forces applied to the right and left wheels of the robot, respectively (assuming the wheels move without slipping, τ_R and τ_L are equal to the torque of the respective motors), d is the distance from the center of mass of the robot to the axis passing through the centers of the robot's wheels, and $m_c$is the mass of the mobile robot without the wheels and the motors attached to them, $m = m_c + 2 \cdot m_w$ is the total mass of the robot (m_w is the mass of the wheel with the attached), $I = I_c + m_c d^2 + 2m_w L^2 + 2I_m$ is the moment of inertia of the mobile robot about the vertical axis passing through point A (the center of the wheel axle of the robot), and I_c is the moment of inertia of the robot (without the wheels and motors) about the vertical axis through the center of mass, I_m is the moment of inertia of the wheel about the vertical axis through the center of the wheel, and I_w is the moment of inertia of the robot's wheel about the horizontal axis passing through the center of the wheel. Next, by considering the kinematic equations (6) and (7), we can derive the dynamic equations in the form of (10):

$$\begin{cases} \ddot{\varphi}_R = f_1(\dot{\varphi}_R, \dot{\varphi}_L . \tau_R, \tau_L) \\ \ddot{\varphi}_L = f_2(\dot{\varphi}_R, \dot{\varphi}_L . \tau_R, \tau_L) \end{cases} \quad (10)$$

We do not provide the analytical expressions for f_1 and f_2 here due to their complexity, but it is worth noting that deriving these expressions is a completely trivial task. The brushed DC motors installed on the robot are described by equations (11):

$$\begin{cases} v_a = R_\alpha + L_\alpha \frac{di_a}{dt} + e_\alpha \\ e_\alpha = K_b \omega_{m\alpha} \\ e_\alpha = K_i i_\alpha \\ \tau = N\tau_m \end{cases}, \quad (11)$$

where i_a is the armature current, (R_α, L_α) (the index α takes values from $\{R, L\}$) is the resistance and inductance of the armature winding respectively, e_α is the back emf, $\omega_{m\alpha}$ is the rotor angular speed , τ_m is the motor torque, (K_i , K_b) are the torque constant and back emf constant respectively, N is the gear ratio, and is τ the output torque applied to the wheel. Since in a DDWMR, the motors are mechanically connected to the robot's wheels via gears, the mechanical equations of the motors' motion are directly related to the mechanical dynamics of the DDWMR. Therefore, for each Direct Current (DC) motor, the following holds (12):

$$\omega_{m\alpha} = N\dot{\varphi}_\alpha \quad (12)$$

From the system (11), taking into account (12), we can derive equation (13):

$$\dot{\tau}_a = \left(\frac{K_t}{L_a}\right)v_a - \left(\frac{R_a K_t}{L_a}\right)\tau_\alpha - \left(\frac{K_t K_b N}{L_a}\right)\dot{\varphi}_\alpha \quad (13)$$

In the four-dimensional phase space $(\dot{\varphi}_R, \dot{\varphi}_L, \tau_R, \tau_L)$ equations (10) and (13) define the system's dynamics in canonical form (14):

$$\dot{x} = f(x, u), \quad (14)$$

where $x(t) = \left(\dot{\varphi}_R(t), \dot{\varphi}_L(t), \tau_R(t), \tau_L(t)\right)$ is the state vector of the system is a $u(t) = \left(u_L(t), u_R(t)\right)$ is the control signal vector.

B. Structure of the Microcontroller Control System for a Two-Wheeled Differential Drive Mobile Robot

We assume that the mobile robot is controlled by a microcontroller or a single-board computer. As a result, we consider a discrete-time control system, with its block diagram shown in Figure 2. The mobile robot is equipped with DC commutator gear motors that have integrated rotary encoders. These encoders produce a set number of pulses per revolution of the gearbox shaft. Processing the encoder data serves two purposes. First, the encoder data can be used to generate discrete signals that represent the relationship between the angular velocity of the robot's wheel rotation and time. Second, access to the encoder data enables the creation of an odometry system for determining the robot's position and orientation (assuming there is no wheel slippage).

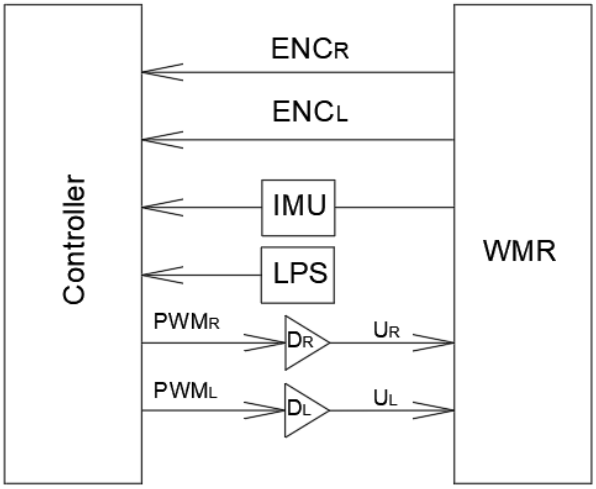


Fig. 2. Block Diagram of the Microcontroller Control System for a Two-Wheeled Mobile Robot

The controller is either a microprocessor board or a single-board computer. D_R and D_L represent the DC commutator motor drivers for the mobile robot. PWM_R and PWM_L are low-voltage pulse-width modulated signals controlling the motor drivers of the right and left wheel, respectively. ENC_R and ENC_L are the signals from the encoders of the robot's right and left motors, IMU (Inertial Measurement Unit) is a Micro-Electro-Mechanical Systems (MEMS) inertial sensor unit, LPS (Local Positioning System) is the local positioning system unit.

C. *Structure of a microcontroller program controlling a mobile robot. Timing organization*

Let T denote the duration of the working cycle of the microcontroller-based control system. At time instants $t_k = k \cdot T$, where k is an integer index, the control signal $\boldsymbol{u}(t)$ is updated. During the working cycle (over the time interval $[t_k, t_{k+1}]$) the controller is engaged in collecting data from sensors. The timing is set by one of the microcontroller's built-in timers. Timing is set by one of the microcontroller's built-in timers. Timer interrupts occur at a fixed time interval τ. During timer interrupt processing, sensors are polled (such as MEMS inertial sensors and local positioning systems) and the acquired readings of discrete signals are stored in corresponding data structures (e.g., circular buffers) in the microcontroller's random access memory. Therefore, the interval τ can be considered as the discretization step for feedback signals. The duration of the working cycle T is chosen so that it contains a whole number M of time intervals τ, i.e. $T = M \cdot T$. The time instants t_k are determined by the overflow of a software-implemented counter modulo M (the counter is incremented at each timer interrupt). Encoder pulses are processed by the microcontroller's built-in capture units. The use of hardware capture units allows measuring the time intervals between encoder pulse edges in the clock cycles of the timer source associated with the respective capture unit. To organize encoder data reading, interrupt processing is performed, the source of which are the engaged capture units. Interrupts are generated upon the arrival of encoder pulse edges (it is obvious that interrupts generated by capture units are not synchronized with timer interrupts). After the working cycle is completed, code implementing a software control unit for managing the mobile robot is executed, i.e., control codes c_k^1 и c_k^2 are calculated based on the readings of discrete sensor signals (feedback signals). Subsequently, after control code calculation, depending on the hardware implementation of the robot's motor drive, the control codes are either loaded into Pulse Width Modulation (PWM) control registers or loaded into digital-to-analog converters (DACs) for voltage controlled drives or transmitted over one of the standard serial communication interfaces (such as UART, SPI, I2C) to digitally controlled drives. Let us $u_1^*(t)$ and $u_2^*(t)$ as step functions, defined by the equations (15).

$$\forall t \in [t_k, t_{k+1}] \quad u_j^*(t) = u_j^k \quad u_j^k = u_{max} \cdot \left[\frac{c_k}{c_{max}}\right], \quad (15)$$

where $j = \{1,2\}$, and $c_{max} = 2^N - 1$. In the last equation, the integer N p is the resolution of the PWM or DAC, used for controlling the mobile robot's motors. Obviously, It is assumed that the values of the control codes c_k^1 и c_k^2 are limited to the value c_{max}, i.e. $\forall k \; c_k^1, c_k^2 \in \{0,1, \dots, c_{max}\}$. Assuredly, the actual control signals should be described by a continuous function $u_j(t)$. We adopt a control signal model of the form (16), where the parameter value $\lambda > 0$ is chosen such that $\frac{1}{\lambda} \ll T$.

$$\forall t \in [t_k, t_{k+1}] \quad u_j(t) = u_j^{k-1} + \\ +\left(1 - e^{-\lambda(t-t_k)}\right)\left(u_j^k - u_j^{k-1}\right). \quad (16)$$

III. Results and Discussion

A. *Structure and principles of operation of a simulator for an autonomous wheeled mobile robot with a microcontroller-based control system*

The simulator is structurally divided into two modules: a hardware environment simulator and a virtual control module. Functionally, the virtual controller represents four user-defined functions – three virtual interrupt handlers with predefined names and signatures, and also a user-defined $Init()$ function, as well as the simulator's API – a set of functions implemented in the hardware environment simulator code, available to the user when writing the code for virtual interrupt handlers. All user-defined functions of the virtual controller are called only from the hardware environment simulator code. The $Init()$ function is called by the hardware environment simulator at the beginning of the simulation process. It is assumed that user-defined variables are initialized and dynamic data structures are created in the code of the $Init()$ function. Furthermore, it is assumed that the call to the $Init()$ function coincides with the beginning of the first working cycle of the simulation process and the user has the ability to set the initial values of the control codes c_0^1 and c_0^2 in the code of the $Init()$ function by calling the simulator API function $SetControlCodes(int\ cc_1, int\ cc_2)$. The actual simulation process is based on sequentially calculating the dependence of the system state vector $\boldsymbol{x}(t) = \left(\dot{\varphi}_R(t), \dot{\varphi}_L(t), \tau_R(t), \tau_L(t)\right)$ on time at time intervals $[t_k, t_{k+1}]$ working cycles of the virtual mobile robot controller). At the beginning of each simulation step, the hardware environment simulator calls a user-defined virtual interrupt handler function from timer $timer\_1\_interrupt\_handler()$. It is assumed that timer 1 interrupts correspond to the end of the current working cycle of the virtual microcontroller. In the interrupt handler, new values of the control codes c_k^1 and c_k^2 are set for the next k – th working cycle ($[t_k, t_{k+1}]$ time interval). Since the values of the control codes c_k^1 and c_k^2 Since the values of the control codes $\boldsymbol{u}(t)$ on the interval $[t_k, t_{k+1}]$ according to equations (16), solving the system dynamics equation ($\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u})$) on the interval $[t_k, t_{k+1}]$ reduces to an initial value problem for an Ordinary differential equation (ODE) system of the form (17):

$$\boldsymbol{x}(t_k) = \boldsymbol{x}_{t_k} \quad \dot{\boldsymbol{x}} = \boldsymbol{g}(\boldsymbol{x}, t), \quad (17)$$

where the initial value $\boldsymbol{x}_{t_k}$ is the value of $\boldsymbol{x}_{pr}(t_k)$ ($\boldsymbol{x}_{pr}(t)$ the dependence of the system state vector $\boldsymbol{x}(t)$ on time on the interval $[t_{k-1}, t_k]$, calculated on the previous simulation step, and the function $\boldsymbol{g}(\boldsymbol{x}, t)$ is defined by equation (18).

$$\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{t}) = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}(\boldsymbol{t})) \quad (18)$$

For numerical solution of the system of ordinary differential equations (4), we used the solve_ivp ODE solver from the SciPy library. Among the input parameters of the solve_ivp there is the *method* parameter, which allows us to choose one of several numerical methods implemented in the library for solving the initial value problem for ODE systems.

We used the Explicit Runge-Kutta method of order 5 (17) [14]. The solution of the ODE system (18) gives discretely defined functions $\dot{\varphi}_R(t)$ and $\dot{\varphi}_L(t)$ on the interval $[t_k, t_{k+1}]$ (dependencies on time of the angular velocities of rotation of the right and left wheels of the mobile robot). First, based on these data, kinematic characteristics such as the

dependence of the magnitude of the linear velocity of the mobile robot on time and the dependence of the angular velocity of the robot on time are restored. Combining these data with the data obtained in previous steps of the simulation process, numerical integration methods are used to calculate the robot's trajectory on the plane. The next step of the simulation process is to perform a sequence of calls to the user-defined virtual interrupt handler function from timer $timer\_2\_interrupt\_handler()$. It is assumed that timer 2 interrupts correspond to the cycles of data collection by the virtual microcontroller program. In the $timer\_2\_interrupt\_handler()$ interrupt handler, data on the current position and orientation of the virtual robot is typically requested. After the procedures described above are completed, the simulation proceeds to the next step for the next working cycle (time interval $[t_{k+1}, t_{k+2}]$).

Although articles [13] and [15] offer valuable perspectives on kinematic and dynamic modeling, their simulation methods are less effective at capturing the intricacies of microcontroller-based control systems. Our approach stands out by providing a more realistic and practical simulation, making it a more robust tool for testing and assessing control algorithms for wheeled mobile robots, especially in educational environments.

IV. Conclusion

The wheeled mobile robot microcontroller system simulator presented in this paper enables testing of program code that implements the software for the robot's control system. This feature sets the simulator apart from those developed using traditional methods, offering exciting possibilities for its application in educational contexts.

This type of simulator is particularly well-suited for organizing laboratory work in courses focused on designing microcontroller control systems. Its straightforward API, accessible when writing virtual controller code, allows users to concentrate on control tasks without being bogged down by the numerous details typical of low-level embedded programming. Furthermore, the simulator facilitates the use of virtual equivalents of control system components that may be unavailable or impractical for real-world experiments.

Acknowledgment

This research has been funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. AP19679327).

References

[1] Ch.. Zhang, Ch. Cen, and J. Huang, "An Overview of Model-Free Adaptive Control for the Wheeled Mobile Robot", World Electric Vehicle Journal, 15, 396, June 2024, pp. 1-25.

[2] A. Fahmi, M. Sulaiman, I. Siradjuddin, A. Risdhayanti, "Fish Swarmed Kalman Filter for State Observer Feedback of Two-Wheeled Mobile Robot Stabilization", International Journal of Robotics and Control Systems, vol. 3, no. 3, 2023, pp. 470-484.

[3] V. T. Ha, T. T. Thuong, L. N. Truc "Trajectory tracking control based on genetic algorithm and proportional integral derivative controller for two-wheel mobile robot", Bulletin of Electrical Engineering and Informatics, vol. 13, no. 4, August 2024, pp. 2348-2357.

[4] A. Waga, A. Ba-ichou, S. Benhlima, A. Bekri, J. Abdouni, "Efficient autonomous navigation for mobile robots using machine learning", IAES International Journal of Artificial Intelligence (IJ-AI), Vol. 13, No. 3, September 2024, pp. 3061-3071.

[5] G. Györök, B. Beszédes, "Fault tolerant power supply systems", Orosz, Gábor Tamás (szerk.) 11th International Symposium on Applied Informatics and Related Areas (AIS 2016). Székesfehérvár, Magyarország: Óbudai Egyetem, 2016, pp. 68-73.

[6] G. Györök, B. Beszédes, "Duplicated Control Unit Based Embedded Faultmasking Systems", IEEE - Szakál, Anikó; IEEE (szerk.) 2017 IEEE 15th International Symposium on Intelligent Systems and Informatics (SISY) : 15th IEEE International Symposium on Intelligent Systems and Informatics. New York, Amerikai Egyesült Államok: IEEE (2017), pp. 283-288.

[7] S. Kim, S. Kwon, "Dynamic Modeling of a Two-Wheeled Inverted Pendulum Balancing Mobile Robot", Int. J. Control Autom., vol. 13, no. 4, 201,5 pp. 926–933.

[8] V. Mudeng, B. Hassanah, Y. T. K. Priyanto, O. Saputra, "Design and Simulation of Two-Wheeled Balancing Mobile Robot with PID Controller", International Journal of Sustainable Transportation Technology, Vol. 3, No. 1, April 2020, pp. 12-319.

[9] H. Xue, Sh.. Lu, Ch. Zhang, "An Adaptive Control Based on Improved GrayWolf Algorithm for Mobile Robots", Appl. Sci. 14, 7092, 2024, pp. 1-24.

[10] M. Faiz, B. Sumantri, and B. S. B. Dewantara, "Double Loop Controller of Four Mecanum-Wheel Automated Guided Vehicle using SMC-PID," in 2023 International Electronics Symposium (IES), IEEE, 2023, pp. 305–310.

[11] G. Györök, B. Beszédes, "Fault-tolerant Software Solutions in Microcontroller Based Systems", Orosz, Gábor Tamás (szerk.) AIS 2017 - 12th International Symposium on Applied Informatics and Related Areas organized in the frame of Hungarian Science Festival 2017 by Óbuda University: Proceedings. Székesfehérvár, Magyarország : Óbudai Egyetem (2017), pp. 7-12.

[12] J. L. Mata-Machuca, L. F. Zarazua, R. Aguilar-L´opez, "Experimental Verification of the Leader-Follower Formation Control of Two Wheeled Mobile Robots with Obstacle Avoidance", IEEE Latin America Transactions, Vol. 19, No. 8, August 2021, pp. 1417-1424.

[13] R. Dhaouadi, A.Abu Hatab, "Dynamic Modelling of Differential-Drive Mobile Robots using Lagrange and Newton-Euler Methodologies: A Unified Framework" Dhaouadi and Hatab, Adv Robot Autom, vol. 2, no. 2, 2013, pp. 1-7.

[14] J. R. Dormand, P. J. Prince, A family of embedded Runge-Kutta formulae, Journal of Computational and Applied Mathematics, vol. 6., no. 1., 1980, pp. 19-26.

[15] M. Zohaib, S. M. Pasha, H. Bushra, Hassan, J. Iqbal "Addressing Collision Avoidance and Nonholonomic Constraints of a Wheeled Robot: Modeling and Simulation", IEEE International Conference on Robotics & Emerging Allied Technologies in Engineering (iCREATE) April 2014, pp. 1-6.

Vintage Nixie Timepiece with Modern Control

Ádám Lantos
Obuda University
Alba Regia Faculty
Székesfehérvár, Hungary
adam.lantos@stud.uni-obuda.hu

Bertalan Beszédes
Obuda University
Alba Regia Faculty
Székesfehérvár, Hungary
beszedes.bertalan@uni-obuda.hu
https://orcid.org/0000-0002-9350-1802

***Abstract*- This paper presents the design of a Nixie tube clock as an educational tool for teaching key concepts in electronics, embedded systems, and programming. By combining vintage Nixie tube displays with modern microcontrollers, students gain practical experience in circuit design, voltage conversation, digital control, and timekeeping algorithms. The project enhances hands-on learning, encouraging active problem-solving while exploring the evolution of display technology. Additionally, the integration of retro components promotes sustainable engineering practices. Initial results suggest that this approach improves student engagement and understanding of electronics and computing fundamentals in engineering education.**

Keywords - Nixie tube, BCD decoder, glow discharge, RTC, real time clock, sustainability, resourcefulness

I. INTRODUCTION

Integrating vintage electronic components, such as Nixie (Numerical Indicator experimental No.1) tubes, with modern technology represents a compelling intersection of history and innovation. This paper examines the potential benefits and challenges of combining vintage and modern electronics, specifically through the creation of a Nixie tube clock, a popular device for demonstrating both fundamental electronics and microcontroller programming.

Nixie tubes, popular in the mid-20th century for displaying numerical data, have a distinctive visual appeal but require higher voltages and analog control circuits. In contrast, modern microcontrollers offer low-power, precise digital control. Combining these technologies requires bridging analog and digital domains, creating an educational opportunity to explore these different approaches.

The primary benefit of combining vintage and modern technologies is the enhanced learning experience. By utilizing a vintage component like the Nixie tube, students are exposed to older technology, which broadens their understanding of how modern electronics evolved. Building a Nixie tube clock requires knowledge of high-voltage circuitry to power the tubes while employing modern microcontrollers to drive them, thus combining analog and digital systems.

From a design perspective, this fusion offers a unique aesthetic that is visually engaging. Nixie tubes have a distinctive glow that modern displays cannot replicate, providing not only a functional timekeeping device but also an artistic, retro-style artifact.

Moreover, the integration encourages sustainability and resourcefulness. Repurposing vintage components like Nixie tubes reduces electronic waste, promoting the reuse of older technologies in novel ways. This concept aligns with the principles of sustainable engineering and can introduce students to eco-friendly design practices.

Despite the educational and aesthetic benefits, several challenges arise when integrating vintage technologies with modern electronics. Technical complexity is one of the primary drawbacks. Nixie tubes require high-voltage drivers, typically 170V, while modern microcontrollers operate at low voltages (3.3V or 5V). This requires careful design of power supplies and voltage conversion circuits, introducing additional layers of complexity that may be challenging for beginners.

There is also the issue of component scarcity. Nixie tubes are no longer in mass production, and finding reliable, functioning tubes can be difficult and costly. This limited availability can make long-term projects less sustainable and more expensive to maintain.

Finally, compatibility issues can arise. Vintage components are not designed to interface with modern microcontrollers directly, necessitating additional hardware like high-voltage transistors or specialized driver, converter ICs. This increases both the cost and the difficulty of the project.

II. THEORETICAL BACKGROUND

The operation of Nixie tubes is based on the gas discharge phenomenon, which takes place in cold-cathode electron tubes filled with neon or other noble gases. The digits inside the tube are represented by individual cathodes, each of which is energized to produce a gas discharge and a bright orange glow around the respective cathode. The operating voltage is typically 170-200V, which is necessary to create the gas discharge [1-3].

During the discharge phenomenon, the gas field around the cathode is ionized, and the liberated electrons collide with the gas molecules and emit photons [4, 5]. This creates the characteristic light that makes the digits visible. The operating principle of Nixie tubes is based on analog technology, which makes the displayed numbers and signs visually different from modern digital displays, thus giving the devices a unique look.

In Fig. 1., the curve in blue shows the glow discharge section currently under discussion. At point E, the coronal discharge occurs, which can be observed, for example, in Tesla coils. At point F the glow discharge starts and stabilises at point G. In the ascending stage, an "abnormal" discharge occurs, and finally it turns into an arc discharge at point I [6]. It can be clearly seen that the glow discharge we use occurs in the milliampere range in the nx100 V range. Following in the footsteps of this physical phenomenon, a pair of Hungarian brothers, George and Zoltan Haydu in Haydu

Brothers Laboratories, developed the first Nixie tube in the 1930s and used it in the 1950s in different instruments [7].

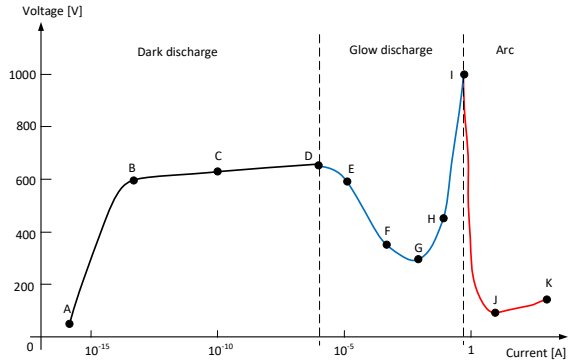


Figure 1: V-I curve of neon gas discharge in constant pressure

The tubes were maintained at pressures ranging from 0.1 to 10 Torr, utilizing a Penning gas mixture composed of neon and argon. The wire-formed elements serve as the cathode (-), while a wire mesh within the tube functions as the common anode (+). When a sufficiently high voltage, approximately 170 V, is applied between the anode and cathode, a yellowish glow is observed around the cathode, indicative of a discharge.

As illustrated in Figure 2, the color of the discharge is predominantly determined by the ionization of neon gas. However, the tube also contains mercury vapor, which emits a blue glow. This addition serves a critical role in preventing "cathode poisoning." Frequent use of the cathode can lead to the formation of oxide layers on adjacent, less-utilized electrodes, potentially impairing performance. The presence of mercury vapor mitigates this oxidation effect. Figure 3 presents the specific Nixie tube model utilized in this study, along with its corresponding pinout.

Figure 2: Fog discharge in Nixie tube

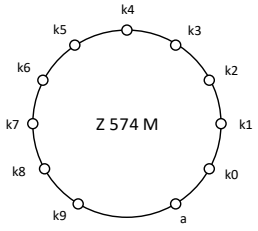


Figure 3: Pinout of the Nixie tube

A. Control of Nixie tubes

The control of Nixie tubes requires a relatively high voltage, which is realized with the help of a suitable power supply and transistors. One of the control solutions is the use of BCD (Binary-Coded Decimal) coded drivers, which receive the numbers in binary form, then decode them and activate the corresponding cathode. This form of control can be clearly integrated with modern microcontroller-based circuits.

To optimize space utilization and eliminate the need for high-voltage transistors (e.g., MPSA42), HVCMOS technology shift registers can be employed. These integrated circuits (e.g., HV509) house all necessary components in a DIL or SO package [8, 9]. Alternatively, general-purpose BCD decoders, such as the HCF4028, can be used in conjunction with transistor arrays. By supplying a low or high signal to a 4-bit input, the corresponding binary number is generated, which the IC uses to activate the output pin corresponding to the given decimal value. Since the anodes are connected to the control circuit, the output is pulled to ground potential.

In 1972, engineers at Texas Instruments developed a decoder specifically designed for controlling cold-cathode display devices, the SN74141. The outputs of this IC are equipped with NPN transistors for switching, and zener diodes are included to prevent negative voltage transitions (Figure 4). This integration reduces circuit complexity and space requirements, as all necessary components are enclosed in one package. Russian equivalents of these ICs, such as the K155ID1, are still available on the market. Given that controlling six ICs would require managing 24 inputs, multiplexing or using a GPIO expander (e.g., MCP23017) is recommended for such configurations.

Enhanced versions of the SN74141 also exist. The SN74142N, for example, includes both a 4-bit decoder and additional features such as a 4-bit latch and a counter. This IC is built using MSI technology and contains four 1-bit master-slave flip-flops. When the $\overline{\text{clear}}$ input is set low, both the counter and flip-flops are reset, resulting in low Q outputs and a high $\overline{Q}_d$ output. While the clear input remains high, the internal counter increments with the rising edge of the clock signal. The $\overline{Q}_d$ output, which is not latched, is routed externally to enable cascading of multiple ICs, thereby facilitating the creation of an n-bit counter. The counter's Q outputs are connected to the data inputs of the latch, and when the strobe input is low, the latch outputs track the counter outputs. If the strobe input is set high, the latch holds the last received value. $\overline{Q}_d$ output is specifically designed for synchronization with subsequent ICs in a cascaded configuration, meaning that the first stage counter can receive new values independent of the latch state [10].

B. Microcontroller program

Modern control of Nixie tube clocks is often solved using modern microcontrollers. Open source platforms are suitable for timing and control tasks. The control program is responsible for the correct operation of the clock, the display of the exact time, and the implementation of additional functions (e.g. temperature measurement, lighting control). The control algorithm typically uses a real-time clock Real-Time Clock, which ensures accurate time tracking even if the

power supply is interrupted. The RTC communicates with the microcontroller via (in this case) I^2C or SPI serial interface, ensuring timing accuracy and easy integration into the system.

The microcontroller updates the time from the RTC, after every significant software event and from interrupt, in every 12-hour interval. When an alarm is set, the current time array is continuously compared to the predefined alarm time array. Upon a match, the alarm function is triggered. In addition to the deactivation function, a snooze feature is available; after five consecutive snooze activations, the alarm is automatically deactivated (see on Fig. 5.). Similar firmware implementations can be found in [11-12]. The microcontroller-based approaches outlined here are also applicable in advanced technical domains [13-14], and they align effectively with educational frameworks in higher education [15-17], particularly with a focus on contemporary cybersecurity education strategies [18-20].

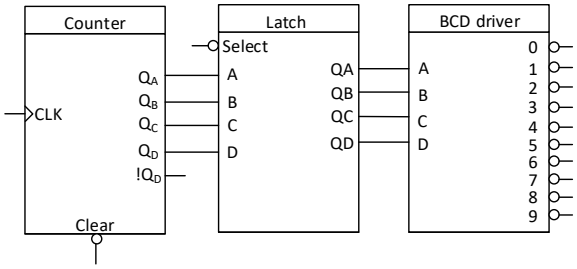


Figure 4: Internal block diagram of the driver IC [11]

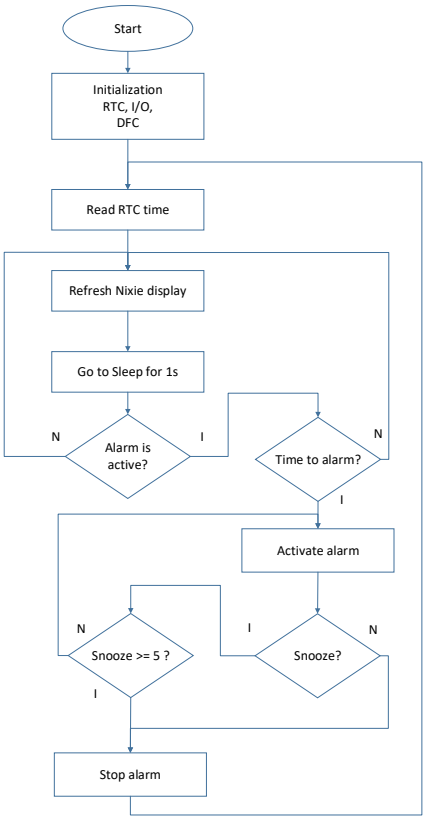


Figure 5: Main block diagram

C. Integration of additional functions

In addition to the time display, Nixie tube clocks can be expanded with other interesting functions that can serve both educational and practical purposes. Nixie tubes can also be used for temperature and humidity display with proper software support and the installation of the necessary sensors, such as DHT22 or DS18B20. With the help of modern microcontrollers, the lighting control of Nixie clocks can also be implemented using a light intensity sensor and, for example, PWM-controlled LEDs to change the intensity of the background lighting. The software control provides the possibility to set different lighting modes, for example the night mode implemented in the project, where the brightness of the display decreases. Furthermore, a Bluetooth or Wi-Fi module can be installed, which enables communication with smartphones, for example setting the time, alarm time or remote control of various lighting schemes.

In this way, students can learn about the integration of wireless communication protocols and embedded systems. Such projects help students understand the integration of vintage technology with modern solutions, which provides useful knowledge in terms of sustainability and upcycling. This approach can also be more sustainable, since the recycling of old technologies can reduce electronic waste by using the extra engineering hours required for implementation. This dual technology perspective allows students to apply the acquired knowledge in unique, tailored projects.

D. Power supply

12 V is supplied from the wall adapter, from which 170 V should be supplied to the tubes and 5 V to the ICs. I did this with a Buck and a Boost converter. The buck converter is homemade, the wiring diagram is shown in Figure 8.

The above equipment (Fig. 7.) is a possible implementation of the 170V circuit. The anodes of the tubes are connected to the output voltage through 15 kΩ protective resistors, the cathodes are connected to the corresponding outputs of the driver ICs. Here, the widely used ZVS step-up DC-DC converter has been used. The converter had been used can be operated from 12V. The buck converter is capable of 12V to 5V conversion, the converter, of which is an LM2576 DC-DC converter, see on Fig. 8.

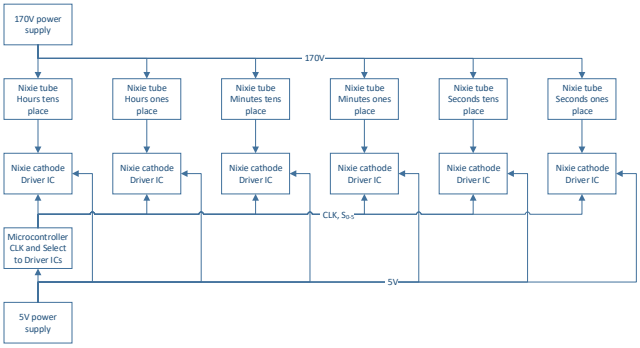


Figure 7: Nixie driver architecture

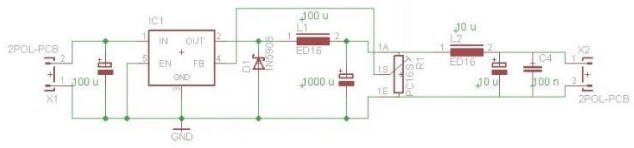


Figure 8: Buck converter circuit diagram

A common issue in switching power supplies is the presence of noise and voltage fluctuations on the output. These arise due to the high-frequency noise introduced by the switching process and the inherent nature of voltage conversion in such circuits. The "spikes" are brief but high-amplitude voltage surges. While these characteristics may not pose problems for some loads, they can negatively impact the performance of sensitive circuits. In this case, since a voltage-sensitive integrated circuit is being powered, it was necessary to address this issue. An effective solution is the use of an LC low-pass filter at the output. This configuration allows low-frequency and DC signals to pass while filtering out high-frequency components. To ensure proper filtering, the cutoff frequency of the Bode plot should be set to a fraction of the switching frequency. Given that the converter's internal oscillator operates at 50 kHz, selecting the cutoff frequency at this range ensures appropriate signal attenuation. In the implementation, a 10 µH inductor and two parallel capacitors, with values of 100 nF and 10 µF respectively, were used for the filter.

III. Practical Realisation

The block diagram (Fig. 9) illustrates the microcontroller, with the ICs arranged in a row beneath it. The lower sections represent the 5 V, 170 V, and 12 V power sources. The RTC and DFC modules communicates with the controller via analog and digital inputs, while the other digital ports function as outputs. One clock signal, other six provide an enable signal, and one delivers a clear signal to the ICs. The ICs receive a 5 V supply, while a 12 V to 170 V boost converter generates sufficient voltage for the tubes. A common ground point has been established within the system using a stub, ensuring that the ICs apply the necessary 170 V ignition voltage to the appropriate tube cathode by pulling the corresponding output to ground.

When using K155ID type ICs, modifications to the microcontroller outputs are required, which involve adding an extra unit. Specifically, a GPIO expander array must be inserted between the microcontroller and the ICs. This allows for the expansion of two microcontroller ports into 16 ports through I^2C communication. By employing two expander lines, the CLK, STRB lines shown in the figure are replaced with SDA, SCL lines to interface with the GPIO. The expander setup provides sufficient connection points to transmit the appropriate binary numeric values to the 4-bit control inputs of the ICs, with each IC receiving four input lines. The ICs will then pass the output signals directly to the tubes. Although this requires more significant changes to the microcontroller's programming, there is library available for the expander that simplifies the integration process.

After the design and modeling of the clock components, the tubes are placed in custom-made sockets mounted on top of the PCB (Fig. 10). In the current configuration, the tubes are connected to the control system using pinheaders.

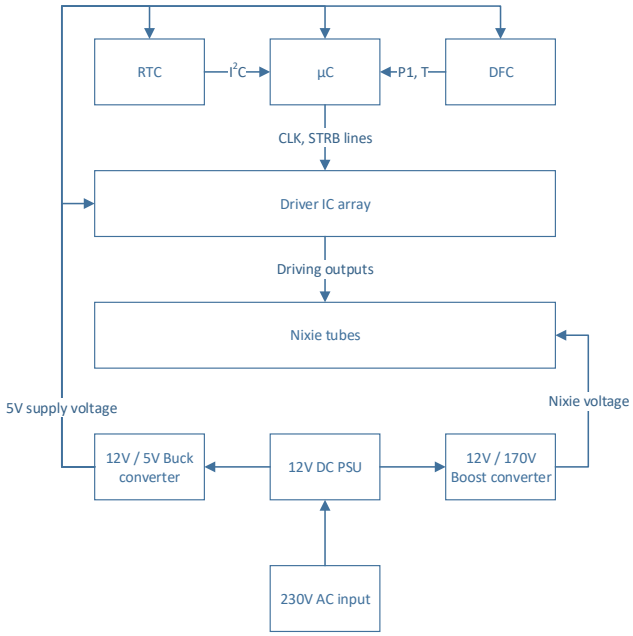


Figure 9: Block diagram of the device

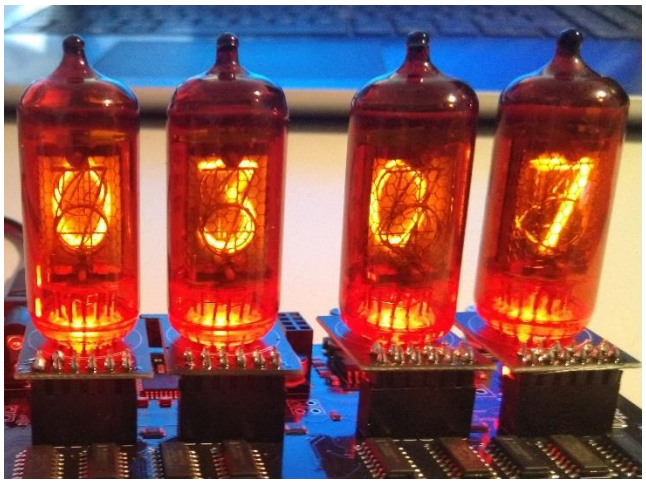


Figure 10: The completed timepiece

Conclusion

The combination of vintage Nixie tube technology with modern microcontroller systems offers both significant educational benefits and technical challenges. While the aesthetic appeal and learning opportunities in analog and digital electronics are substantial, the complexity of high-voltage circuits and the scarcity of vintage components present notable obstacles. Nevertheless, the creation of a Nixie tube clock represents a practical and rewarding project that demonstrates the fruitful intersection of historical and modern technologies in an educational setting. Combining electronic innovation with the intricacies of traditional analog circuits provides future engineers with a solid foundation in challenging industries.

Acknowledgment

The authors would like to thank all the faculty staff and member of Obuda University Alba Regia Faculty, that provide help and assistance throughout the project completion.

REFERENCES

[1] Weston, G.F. Cold Cathode Glow Discharge Tubes. K Barker. 1968 Phys. Bull. 20 237 DOI 10.1088/0031-9112/20/6/006

[2] Lisovskiy, V. A., et al. "Validating the Goldstein-Wehner law for the stratified positive column of dc discharge in an undergraduate laboratory." European journal of physics 33.6 (2012): 1537.

[3] Arumugam, Saravanan, Prince Alex, and Suraj Kumar Sinha. "Effective secondary electron emission coefficient in DC abnormal glow discharge plasmas." Physics of Plasmas 24.11 (2017).

[4] Betti, Maria, and L. Aldave de las Heras. "Glow discharge mass spectrometry in nuclear research." Spectroscopy Europe 15.3 (2003): 15-24.

[5] Mathew, Prijil, et al. "Experimental verification of modified Paschen's law in DC glow discharge argon plasma." AIP Advances 9.2 (2019).

[6] C. F. Gallo, "Coronas and Gas Discharges in Electrophotography: A Review," in *IEEE Transactions on Industry Applications*, vol. IA-11, no. 6, pp. 739-748, Nov. 1975, doi: 10.1109/TIA.1975.349370.

[7] Dieter Waechter. Dieter's Nixie Tube Data Archive https://www.tube-tester.com/sites/nixie/dat_arch/RFT_book_05.pdf (2024.09.09.)

[8] Motorola Inc. Semiconductor Technical Data, High Voltage Transistors, NPN Silicon, MPSA42/D Datasheet

[9] Supertex Inc. 16-Channel Serial to Parallel Converter with High Voltage Backplane Driver and Push-Pull Outputs, HV509 datasheet

[10] Texas Instruments - The TTL Data Book for Design Engineers. Second edition. 1981. LCC4112 74062-116 –AI

[11] Alexander Baklanov, Svetlana Grigoryeva, György Györök. Control of LED Lighting Equipment with Robustness Elements, Acta Polytechnica Hungarica (1785-8860 1785-8860): 15 3 pp 105-119 (2016)

[12] G. Gyorok and M. Mako, "Configuration of EEG input-unit by electric circuit evolution," *2005 IEEE International Conference on Intelligent Engineering Systems, 2005. INES '05.*, Spain, 2005, pp. 255-258, doi: 10.1109/INES.2005.1555168.

[13] Aizhan Zhaparova, Dimitry Titov, Alexander Y. Baklanov, *Gyorgy* Gyorok. Study of the Effectiveness of Switching-on LED Illumination Devices and the Use of Low Voltage System in Lighting. Acta Polytechnica Hungarica (1785-8860 1785-8860): 12 5 pp 71-80 (2015)

[14] Mosavi Amirhosein, Bertalan Beszedes, Imre Felde, Nadai Laszlo, Gorji Nima E. Electrical characterization of CIGS thin-film solar cells by two- and four-wire probe technique. MODERN PHYSICS LETTERS B 2020 p. 2050102 , 16 p. (2020)

[15] Tolner Nikoletta, Pogátsnik Monika, Módné Takács Judit. A mesterséges intelligencia szerepe az online vizsgáztatásban. ISKOLAKULTÚRA: PEDAGÓGUSOK SZAKMAI-TUDOMÁNYOS FOLYÓIRATA 33 : 10 pp. 39-55. , 17 p. (2023)

[16] Judit MÓDNÉ TAKÁCS, Nikolett TOLNER, Monika POGÁTSNIK. University teachers' perceptions of AI integration: Insights from a qualitative focus group study. OPUS ET EDUCATIO: MUNKA ÉS NEVELÉS 10 : 3 pp. 230-235. , 6 p. (2023)

[17] Nikoletta Tolner, Judit Módné Takács, András Dávid, Monika Pogatsnik. Online examination from the point of view of teachers. In: Petőné Csuka, Ildikó; Simon, Gyula (szerk.). AIS 2022 - 17th International Symposium on Applied Informatics and Related Areas – Proceedings. Székesfehérvár, Magyarország : Óbudai Egyetem (2022) 180 p. pp. 169-173. , 5 p.

[18] Judit Módné Takács, Monika Pogátsnik. A Comprehensive Study on Cybersecurity Awareness: Adaptation and Validation of a Questionnaire in Hungarian Higher Technical Education. ACTA POLYTECHNICA HUNGARICA 21 : 10 pp. 533-552. 20 p. (2024)

[19] Judit Módné Takács, Monika Pogátsnik. The Presence of Cybersecurity Competencies in the Engineering Education of Generation Z. ACTA POLYTECHNICA HUNGARICA 21 : 6 pp. 107-127. 21 p. (2024)

[20] Judit Módné Takács, Monika Pogatsnik. A systematic review of Human Aspects in Industry 4.0 and 5.0: Cybersecurity Awareness and Soft Skills. In: Szakál, Anikó IEEE 27th International Conference on Intelligent Engineering Systems 2023 (INES 2023). Budapest, Magyarország : IEEE Hungary Section (2023) pp. 33-39. 7 p.

POSTER SESSION I.

Information and Computer Systems

Differential-Geometric Methods for Modeling and Control of a Robot Manipulator

Gaukhar Nazenova
School of Digital Technologies and Artificial Intelligence
D Serikbayev East Kazakhstan Technical University
Ust-Kamenogorsk, Kazakhstan
nazenova.g@edu.ektu.kz

Albina Kadyroldina
School of Digital Technologies and Artificial Intelligence
D Serikbayev East Kazakhstan Technical University
Ust-Kamenogorsk, Kazakhstan
akadyroldina@gmail.com

Alexander Krasavin
School of Digital Technologies and Artificial Intelligence
D Serikbayev East Kazakhstan Technical University
Ust-Kamenogorsk, Kazakhstan
akrassavin@ektu.kz

Darya Alontseva
School of Digital Technologies and Artificial Intelligence
D Serikbayev East Kazakhstan Technical University
Ust-Kamenogorsk, Kazakhstan
dalontseva@ektu.kz

Assel Kussaiyn-Murat
School of Digital Technologies and Artificial Intelligence
D Serikbayev East Kazakhstan Technical University
Ust-Kamenogorsk, Kazakhstan
akussaynmurat@edu.ektu.kz

***Abstract* — Nowadays, industrial robots are widely used in modern manufacturing fields such as coating spraying, welding, packing, pick and place, etc. All such industrial applications require the robot end effector to achieve precise positioning and fast response under external disturbances and changing loading conditions. Therefore, high-speed and high-precision robot control is particularly important and has been an area of intensive research. Some of the most promising control methods for industrial robot manipulators are based on the inverse dynamic model. One such method, computational torque control, is an effective way to improve the motion control performance of robot. However, calculating the joint torque is quite complex, and due to the uncertainty of the parameters, it is difficult to obtain an accurate inverse dynamic model of the robot to generate torque. This paper proposes an efficient inverse dynamic model of an industrial robot based on Lie algebra and discusses the applications of this model to the adaptive control of a manipulator.**

Keywords — robot manipulator control, computed torque control, inverse dynamic model, Lie algebra.

I. Introduction

Robot manipulators are widely used in tasks such as pick and place, painting, welding, grinding, milling, and drilling. These operations demand high positioning accuracy, fast response, and stability under external disturbances from the robot [1]. Integrating advanced control methodologies into robotics not only enhances precision and adaptability but also creates valuable educational opportunities, as seen in various engineering domains [2]. G. Györök and B. Beszédes described in their papers that to ensure the reliability and stability of manipulator control, control systems can employ redundant architectures [3,4]. Therefore, high-speed and precise control is a critical aspect of robotics, and various algorithms have been developed to address this, including optimal control, sliding mode control, and computed torque control. The latter is the most effective method for improving dynamic performance and motion accuracy, as it compensates for nonlinearities through feedback and allows the use of linear control methods [5].

However, calculating the inverse dynamics of a robot is a complex task, and accurately determining model parameters is challenging, which limits the application of this approach. Traditional dynamic modeling methods, such as the Newton-Euler and Lagrange methods, are also computationally intensive, especially for robots with flexible joints [6]. Recently, increasing attention has been given to the use of Lie groups and Lie algebras for modeling the kinematics and dynamics of robots [7,8]. These approaches enable the creation of more accurate dynamic models that can be applied in computed torque control [9]. However, due to the numerous parameters, nonlinearities, and time-varying dynamics of robots, developing precise models remains a complex task. In robotics, orthonormal coordinate systems are traditionally used to describe kinematics and dynamics [10], and the Denavit-Hartenberg (DH) convention is often applied to solve kinematic analysis problems. In the work by Zakey et al. [11], it is noted that when using this convention, a specific set of rules must be followed to set up coordinate systems and express translational and rotational relationships between them. While this method is widely used, it has several drawbacks. First, there are different versions for defining coordinate systems, each requiring different sets of rules. For example, parallel axes may require special treatment since their description can be ambiguous. Second, a large number of coordinate systems are needed, which complicates working with robots that have a high number of degrees of freedom (DOF). Third, kinematics and dynamics are tied to fixed coordinate systems, making it difficult when reconfiguring the robot. Salman et al. [12] point out that when a robot is reconfigured, DH parameters must be reassigned, and different types of joints must be considered. Finally, the DH convention limits the choice of stationary and body-fixed frames, which makes it challenging for algorithms that describe the movement of multiple points on a robot, such as whole-body control [13].

In contrast, differential geometry offers a more abstract approach by describing systems as manifolds [14]. Alhousani et al. [15] emphasize that this approach significantly reduces the number of coordinate systems required to describe robot kinematics and dynamics. The advantages of geometric methods have been thoroughly discussed in several papers [15,16]. Although studies comparing traditional and

geometric approaches often focus on computational and algorithmic aspects [17-19], in-depth discussions of their conceptual and practical differences, as noted in [20], are rare.

Thus, computed torque control (CTC) is the most direct and effective method for improving the performance of robot motion control. However, calculating the torque at the joints is quite complex, and due to parameter uncertainties, it is difficult to obtain an accurate inverse dynamic model of the robot for torque generation. This paper examines the geometric model of a serial robot arm and proposes an inverse dynamic model based on this robot model. The paper also describes the CTC control of the robot arm and discusses the possibility of developing an adaptive control system for the manipulator as a modification of the CTC method.

II. Lie algebra formulation of manipulator kinematics and dynamic

A. Lie groupes as smooth mainfolds

In general, when we talk about a group, we mean a certain set of transformations, where the group multiplication operation corresponds to the composition of these transformations. Naturally, for a set with a given multiplication operation to be considered a group, additional conditions must be satisfied, known as the group axioms: for example, the existence of an inverse element and the associativity of multiplication. Accordingly, Lie groups represent groups of continuous transformations. Formally, a Lie group can be defined as follows:

Definition 1. A Lie group is a group G, equipped with a manifold structure such that the group operations

$$Mult: G \times G \to G, \qquad (g_1, g_2) \to g_1 g_2$$
$$Inv: G \to G, \qquad g \to g^{-1}$$

are smooth.

An important example for the applications we are interested in is the group of rotational transformations $SO(3)$. Lie group $SO(3)$ sometimes referred as "group of rotation matrixes R". That means that the set of square orthogonal matrices R, i.e. matrices satisfying the condition $R^T R = E$, with a dimension of 3×3, forms a group under the operation of matrix multiplication. However, the matrix representation of the rotation group is just one of the possible representations. A geometric interpretation of the group $SO(3)$ as a smooth three-dimensional surface embedded in four-dimensional space is more fruitful from a theoretical perspective, as this interpretation allows the application of techniques from smooth manifold theory to this Lie group. In the following discussion, we will illustrate the general concepts of Lie theory using this group of transformations as an example.

Let us remind some definitions related to tangent vector fields on manifolds, which will play a significant role in the subsequent discussion. For arbitrary manifold M we will denote $\mathfrak{X}(M)$ as the set of all smooth tangent vector fields on M. Let $X \in \mathfrak{X}(M)$ a smooth tangent vector field, An integral curve of X at $p \in M$ is a smooth curve $\gamma: I \to M$, where $I \subset \mathbb{R}$ is an interval, such that $\gamma(0) = p$ and for all $t \in I$ equation (1) holds.

$$\dot{\boldsymbol{\gamma}}(t) = X_{\gamma(t)} \tag{1}$$

In local coordinates, equation (1) represents an ODE. Here, we will not focus on the conditions for the existence and uniqueness of the solution to (1), assuming they are satisfied in the following discussion. Vector field X is called *complete* if for all $p \in M$ it is maximum integral curve with $\gamma(0) = p$ is defined for all $t \in \mathbb{R}$. Every complete vector filed X generate a map $\Phi^X: \mathbb{R} \times M \to M \ \ (t, m) \to \Phi_t^X(m)$ called *flow*. For given $X \in \mathfrak{X}(M)$ and arbitrary point $p \in M$ flow of X is defined as (2).

$$\Phi_{-t}^X(p) = \boldsymbol{\gamma}(t), \tag{2}$$

where $\boldsymbol{\gamma}(t)$ is integral curve of X such that $\boldsymbol{\gamma}(0) = p$. Then, for a given function $f: M \to \mathbb{R}$, the covariant derivative $D_X(f)$ of f with respect to X is a function g, which intuitively measures the change in f with respect to the flow under X. Formally the covariant derivative $D_X(f)$ of f is defined by formula (3).

$$q(p) = D_X(f)(p) = \lim_{t \to 0} \frac{f\left(\Phi_{-t}^X(p)\right) - f(p)}{t} \tag{3}$$

By definition q is a mapping $q: C^\infty(M) \to C^\infty(M)$. Therefore, we can consider arbitrary vector field $X \in \mathfrak{X}(M)$ acting on smooth scalar functions defined on M. For given $X, Y \in \mathfrak{X}(M)$ *the composition of these two operators will be denoted as* $X \circ Y$. It is a general fact that the commutator of derivations of an algebra is again a derivation. Thus, $\mathfrak{X}(M)$ is a Lie algebra for the bracket (4).

$$[X, Y] = X \circ Y - Y \circ X \tag{4}$$

Let us remind, that a Lie algebra is just a vector space V, equipped with a map $[\cdot, \cdot] : \mathrm{V} \times \mathrm{V} \to \mathrm{V}$ satisfying the following properties for all vectors u, v, w in V.

1. Bilinearity: [au + bv, w] = a[u, w] + b[v, w]
2. Antisymmetry: [v, w] = −[w, v],
3. Jacobi Identity: [u, [v, w]] + [w, [u, v]] + [v, [w, u]] = 0.

An example of a Lie algebra can be the three-dimensional Euclidean space, where the Lie bracket is defined by the operation of the cross product of two vectors.

B. Lie algebra of Lie grope

As will be shown below, for any Lie group G on the tangent space $T_e(G)$ one can define Lie brackets. Thus, the linear space $T_e(G)$ can be considered as a Lie algebra (this construction is called the Lie algebra of Lie group G). Oddly enough, the Lie algebra almost completely determines the Lie group itself. One can say that an arbitrarily small neighborhood of the identity of a continuous group contains almost complete information about the structure of the entire group as a whole. In particular, each vector $T_e(G)$ can be assigned an element of the group G. In our opinion, the easiest way to obtain the corresponding results is to first consider the concepts of left-invariant vector fields. Let G be a Lie group. Denote by $\mathfrak{g} = T_e(G)$ the tangent space to the group unit. The maps $L_g: G \to G$ and $R_g: G \to G$ defined by $L_g(h) = gh$ and $R_g(h) = hg$, defined for any $g \in G$, are called left- and right-translation. It's easy to proof that both L_g and R_g are diffeomorphisms of G. The differential $DL_g: T_h(G) \to T_{gh}(G)$ gives a natural identification of tangent spaces. A vector field $X \in \mathfrak{X}(G)$ is called left-invariant if for any $g \in G \quad DL_g X = X \circ L_g$ i.e. $DL_g(h)X(h) = X(gh)$. It's almost obvious that set of all left-invariant vector fields $\mathfrak{X}^L(G)$ may be considered as a linear space. But the fact that

for two for two arbitrary $,Y \in \mathfrak{X}^L(G)$ $[X,Y] \in \mathfrak{X}^L(G)$ is difficult to call obvious, although its proof is not difficult [21]. Therefore, $\mathfrak{X}^L(G)$ is Lie algebra. For $\xi \in \mathfrak{g}$ we denote by $\xi^L \in \mathfrak{X}^L(G)$ the unique left-invariant vector field such that $\xi^L|_e = \xi$. Similarly, ξ^R denotes the unique right-invariant vector field such that $\xi^R|_e = \xi$. Since each vector $\xi \in \mathfrak{g}$ uniquely corresponds to a left-invariant vector field $\xi^L \in \mathfrak{X}^L(G)$Lie brackets in $\mathfrak{X}^L(G)$ induce Lie brackets acting on vectors of $\mathfrak{g}$. Now we can give the following definition:

Definition 2. The Lie algebra of a Lie group G is the vector space $\mathfrak{g} = T_e(G)$ equipped with the unique Lie bracket such that the map $\mathfrak{X}(G)^L \to G$, $X \to X_e$ is an isomorphism of Lie algebras.

Of course, it is also possible to use right-invariant vector fields to define a Lie algebra structure; it turns out that the resulting bracket is obtained simply by a sign change. The integral curves of left-invariant vector fields possess a remarkable property. If we denote $\boldsymbol{\gamma}^{\xi}$ as the integral curve of the left-invariant vector field $\xi^L \in \mathfrak{X}^L(G)$ (assuming that $\dot{\gamma}^{\xi}(0) = \xi$) than for any t_1, t_2 equality (5) holds true.

$$\boldsymbol{\gamma}^{\xi}(t_1 + t_2) = \boldsymbol{\gamma}^{\xi}(t_1)\boldsymbol{\gamma}^{\xi}(t_2) \tag{5}$$

In fact, this equality means that the elements of the group G, lying on the integral curve $\boldsymbol{\gamma}^{\xi}$ form a subgroup of G. The elements of this subgroup, located in a small neighborhood of the unit of G, generate the entire "larger" subgroup. This fact forms the basis for the following construction, which is of fundamental importance in the theory of Lie groups.

Definition 3. The exponential map for the Lie group G is the smooth map defined by (6).

$$exp: \mathfrak{g} \to G\ , \xi \to \boldsymbol{\gamma}^{\xi}, \tag{6}$$

where $\boldsymbol{\gamma}^{\xi}$ is integral curve of left-invariant vector field $\xi^L \in \mathfrak{X}^L(G)$ (and also $\boldsymbol{\gamma}^{\xi}$ is the 1-parameter subgroup) with $\dot{\boldsymbol{\gamma}}^{\xi}(0) = \xi$.

Now let's consider the application of the theory to a specific example of the rotation group $SO(3)$. Note that for any rotation matrix R holds equality $R^T(t)R(t) = I$. By differentiating the last equality with respect to time, we obtain $\dot{R}^T R + R^T \dot{R} = 0$. This expression can be rearranged to (7).

$$R^T \dot{R} = -\left(R^T \dot{R}\right)^T \tag{7}$$

This expression reveals that $R^T \dot{R}$ is a skew-symmetric matrix $[\omega]_\times$ and have the form (8).

$$[\omega]_\times = \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix} \tag{8}$$

When R = I we have

$$\dot{R} = [\omega]_\times$$

that is, set of a skew-symmetric matrices $[\omega]_\times$ is in the Lie algebra of $SO(3)$, which we name $so(3)$. Since $[\omega]_\times \in so(3)$ has 3 degrees of freedom, the dimension of $so(3)$ (and as a consequence, the dimension $SO(3)$) is 3. The Lie algebra is a vector space whose elements can be decomposed into (9).

$$[\omega]_\times = \omega_x \boldsymbol{E}_x + \omega_y \boldsymbol{E}_y + \omega_z \boldsymbol{E}_z, \tag{9}$$

where $\boldsymbol{E}_x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$, $\boldsymbol{E}_y = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}$, $\boldsymbol{E}_z = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

These matrices, referred to as the *generators* of $so(3)$, when considered as vectors, form a basis of the vector space $so(3)$. The quantities ω_x, ω_y and ω_z represent the components of the angular velocity vector of the rigid body $\boldsymbol{\omega} = (\omega_x, \omega_y, \omega_z) \in \mathbb{R}^3$. Equation () states that the tangent space so(3) is isomorphic to $\mathbb{R}^3$. We use so called *Hat* operator, defined by (), to pass vectors of $\mathbb{R}^3$ to 3 dimensional linear space of skew-symmetric matrices.

$$Hat: \mathbb{R}^3 \to so(3); \quad \boldsymbol{\omega} \to \widehat{\omega} = [\omega]_\times \tag{10}$$

C. Rigid body motion and special Euclidian Group SE(3)

We will consider the motion of a rigid body in a fixed world Cartesian coordinate system W with origin O. We will assume that a Cartesian coordinate system B, with origin O' s rigidly attached to the rigid body. Generally, rigid body transformations are described by the pairs (R, t) with R describing the rotation, and $\boldsymbol{t}$ describing the translation component. More specifically, R is the transition matrix between the bases W and B (since both W и B are Cartesian bases, the matrix R will be a rotation matrix, i.e. $R \in SO(3)$) and $t \in \mathbb{R}^3$ is just a offset vector OO'. Suppose that there are two joints, the second joint is attached to the first one, while the rigid body attached to the second one with position p. Let the rotation of the first joint be defined by R_1, and the translation generated because of the offset of the rotation axis by t_1, while the rotation of the second joint by R_2, and the translation generated because of the offset of the second joint axis by t_2. Then the position $p'$of the rigid body after the rotations is acquired by application first the transformation (R_2, t_2) and then (R_1, t_1) resulting in (11).

$$p' = R_1 R_2 p + R_1 t_2 + t_1 \tag{11}$$

The composition rule of the offset rotations and rigid body transformations (R_1, t_1) and (R_2, t_2) is (12).

$$(R_1, t_1) \circ (R_2, t_2) = (R_1 R_2, R_1 t_2 + t_1) \tag{12}$$

This model, however, is based on three-dimensional vector geometry, which cannot naturally account for offsets. To address this limitation, the model is extended so that each point and vector is represented by a four-dimensional vector. A four-dimensional vector represents a geometric point when its last coordinate is one, and a geometric vector when its last coordinate is zero. It is easy to see that the difference between points in this representation results in a vector, though the sum of points or the difference between more than two points is undefined. On the other hand, the sum and difference of any number of vectors are well-defined.

The interpretation of the four-dimensional vector is as follows: the first three coordinates represent a geometric vector, and if the fourth coordinate is one, this vector is associated with a base point at the origin of the reference frame, meaning it represents the position of a geometric point. If the fourth coordinate is zero, the vector has no fixed base point and can be freely translated in space, making it a geometric vector. This representation effectively handles geometric objects with offsets and is called the homogeneous

representation, where the coordinates in this four-dimensional space are referred to as homogeneous coordinates. The application of offset rotation in homogeneous coordinates is defined as (13).

$$\begin{pmatrix} p' \\ 1 \end{pmatrix} = \begin{pmatrix} R & 0 \\ t & 1 \end{pmatrix} \begin{pmatrix} p \\ 1 \end{pmatrix} \quad (13)$$

Thus, when using homogeneous coordinates, rigid body motions can be represented by 4×4, the structure of which is described by expression (14).

$$SE(3) := \left\{ g := \begin{pmatrix} R & 0 \\ t & 1 \end{pmatrix} \; R \in SO(3), p \in \mathbb{R}^3 \right\} \quad (14)$$

It is straightforward to confirm that the set of matrices $SE(3)$ meets all the requirements of a group. It is sufficient to demonstrate that the inverse of any element in $SE(3)$ exists and is also a member of $SE(3)$:

$$\begin{pmatrix} R & 0 \\ t & 1 \end{pmatrix}^{-1} = \begin{bmatrix} R^T & -R^T p \\ 0 & 1 \end{bmatrix} \in SE(3) \quad (15)$$

Analogous to how we constructed the matrix representation of the Lie algebra $so(3)$, for an arbitrary $g(t) \in SE(3)$ let us consider the expression (16).

$$\dot{g}(t)g^{-1}(t) = \begin{bmatrix} R\dot{(t)}R^T(t) & \dot{p}(t) - \dot{R}(t)R^T(t)p(t) \\ 0 & 0 \end{bmatrix} (16)$$

Note that $\hat{\omega}(t) = R\dot{(t)}R^T(t)$ and $\dot{p}(t) - \hat{\omega}(t)p(t) = \boldsymbol{v}(t)$ are the world angular and linear velocities of the point in the body that corresponds with the origin of the world frame. Accordingly, the expression (16) can be rewritten in the form of (17).

$$\dot{g}(t)g^{-1}(t) = \begin{bmatrix} \hat{\omega}(t) & \boldsymbol{v}(t) \\ 0 & 0 \end{bmatrix} \quad (17)$$

Consequently, the Lie algebra of the Lie group SE(3) is defined as the set given by (18).

$$se(3) := \left\{ \begin{bmatrix} \hat{\omega}(t) & \boldsymbol{v}(t) \\ 0 & 0 \end{bmatrix} \; \hat{\omega} \in so(3), \boldsymbol{v} \in \mathbb{R}^3 \right\} \quad (18)$$

Elements of $se(3)$ is called *twists*. Each twist $\hat{\xi} = \begin{bmatrix} \hat{\omega}(t) & \boldsymbol{v}(t) \\ 0 & 0 \end{bmatrix}$ can be regarded as result of application hat operator to vector $\xi = (\hat{\omega}, v) \in \mathbb{R}^6$. A tangent vector $\hat{\xi} \in so(3)$ approximates $g(t) \in SE(3)$ locally (19).

$$g(t + \Delta t) \approx \left(I + \hat{\xi}\Delta t\right) g(t) \quad (19)$$

Accordingly, for constant twist $\hat{\xi} \in so(3)$ and $g(t) \in SE(3)$ $\dot{g}(t) = \hat{\xi} g(t)$. Taking into account that $g(0) = I$ we obtain at an important formula for the application of Lie theory to manipulator kinematics (20).

$$g(t) = exp\left(\hat{\xi}t\right) \quad (20)$$

III. Application of Lie Theory Based Model of Rigid Link Serial Manipulator to Manipulator Inverse Dynamic Problem

Utilizing the Lie group structure of rigid body motions allows the kinematics of an open-chain robot to be represented as a product of matrix exponentials. If the revolve joint between adjacent rigid bodies is regarded as a twist $\hat{\boldsymbol{\xi}}$, then the position and attitude of the link i relative to the base coordinate system can be expressed as (21).

$$g_i = e^{\hat{\xi}_1 \theta_1} e^{\hat{\xi}_2 \theta_2} \ldots e^{\hat{\xi}_i \theta_i} M_{i0}, \quad (21)$$

where $\theta_1, \theta_2, \ldots, \theta_n$ are joint variables $M_{i0} \in SO(3)$ is the position and attitude of the i-th link relative to the base coordinate system when the robot is at zero position. Through forward recursion (for$i = 1,2, \ldots, n$) the six dimensional generalized velocity and acceleration can be calculated as follows:

$$\boldsymbol{V}_i = g_i^{-1} \frac{d(g_i)}{dt} = Ad_{g_{i-1,i}^{-1}}(\boldsymbol{V}_{i-1}) + \xi_i \dot{\theta}_i \quad (22)$$

$$\boldsymbol{A}_i = \frac{d(\boldsymbol{V}_i)}{dt} = Ad_{g_{i-1,i}^{-1}}(\boldsymbol{A}_{i-1}) - ad_{\xi_i \theta_i}(\boldsymbol{V}_i) + \xi_i \ddot{\theta}_i \quad (23)$$

The generalized forces $\boldsymbol{F}_i$ of i – th link in link frame are determined in recursive manner by consequential application of equations () (that means $i = n, n-1, \ldots$).

$$\boldsymbol{F}_i = \boldsymbol{E}_i A_i - ad_{\boldsymbol{V}_i}^T(\boldsymbol{E}_i \boldsymbol{V}_i) + Ad_{g_{i,i+1}^{-1}}^T(\boldsymbol{F}_{i+1}), \quad (24)$$

where $\boldsymbol{E}_i \in \mathbb{R}^{6\times6}$ is the mass-inertia matrix of the i-th link, defined as (25).

$$\boldsymbol{E}_i = \begin{bmatrix} \boldsymbol{I}_i - m_i \hat{\boldsymbol{r}}_i^2 & m_i \hat{\boldsymbol{r}}_i \\ -m_i \hat{\boldsymbol{r}}_i & m_i \boldsymbol{I} \end{bmatrix}, \quad (25)$$

where I is a 3×3 identity matrix, and $\hat{\boldsymbol{r}}_i$ is the skew-symmetric matrix of the position vector $\boldsymbol{r}_i \in \mathbb{R}^3$ in link frame $\{i\}$. So, the mass-inertial parameters of the i-th link in equation (25) are contained in matrix $\boldsymbol{E}_i$.

A. *Computed Torque Control. Prospects for applying the derived inverse dynamic model in computed torque-based manipulator control methods*

First and foremost, it is important to note that there are currently many variations of computed torque-based control methods. To start, we will briefly outline the core concepts of the CTC (Computed Torque Control) method, which are common across its various modifications. The CTC method was originally developed for trajectory control of manipulators. From this point on, we will assume that the desired trajectory of the manipulator is given by the function $\boldsymbol{q}_d(t)$, where $\boldsymbol{q} \in \mathbb{R}^N$ is the vector of joint variables of n–link robot arm. Robot arm dynamics of the robot arm are described by the equation (26).

$$M(q)\ddot{q} + V(q, \dot{q}) + \dot{G}(q) + \tau_d = \tau, \quad (26)$$

where $M(q)$ is $n \times n$ mass matrix of the manipulator, $V(q, \dot{q})$) is $n \times 1$ vector of centrifugal and Coriolis terms, $G(q)$ is $n \times 1$ vector of gravity term, and τ_d is disturbance.

For the sake of convenience, we will use a simplified form of equation (27), provided below.

$$M(q)\ddot{q} + N(q, \dot{q}) + \tau_d = \tau \quad (27)$$

We define the tracking error $e(t)$ by the equation (28):

$$e(t) = q_d(t) - q(t) \quad (28)$$

By differentiating the last expression twice with respect to time, we obtain equation (29):

$$\ddot{e}(t) = \ddot{q}_d(t) + \ddot{q}(t) \quad (29)$$

Expressing $\ddot{q}$ from equation (29) and substituting the result into (30), we obtain:

$$\ddot{e}(t) = \ddot{q}_d + M^{-1}(N + \tau_d - \tau) \quad (30)$$

We define the control input u as (31), and the disturbance function w as (32):

$$u = \ddot{q}_d + M^{-1}(\mathrm{N} - \tau) \quad (31)$$

$$\mathrm{w} = \mathrm{M}^{-1}\tau_{\mathrm{d}} \tag{32}$$

Now we may define state $\boldsymbol{x}(t) \in \mathbb{R}^{2n}$ by equation (33):

$$\boldsymbol{x}(t) = \begin{pmatrix} e(t) \\ \dot{e}(t) \end{pmatrix} \tag{33}$$

and write the equation tracking error dynamics in the form (34):

$$\frac{d}{dt}\boldsymbol{x} = A\boldsymbol{x} + B\boldsymbol{u} + C\boldsymbol{w}, \tag{34}$$

where $A = \begin{bmatrix} 0 & I \\ 0 & 0 \end{bmatrix}$,$B = \begin{bmatrix} 0 \\ I \end{bmatrix}$ and $C = \begin{bmatrix} 0 \\ I \end{bmatrix}$. The feedback linearization equation (35) can be inverted.

$$\tau = M(\ddot{q}_d - u) + N \tag{35}$$

The final equation represents the computed torque law. This means the following: if we choose the control input $u(t)$ that stabilizes the tracking error dynamics equation (34), then the nonlinear control input t , calculated according to (35), will ensure that the robot arm follows the desired trajectory. In essence, the nonlinear transformation (31) simplifies the problem of controlling a complex nonlinear system into a straightforward task of controlling a linear system, consisting of n independent linear subsystems. Clearly, the computed torque depends on the inversion of the robot's dynamics. For this reason, the CTC method is sometimes referred to as *inverse dynamics control*. As seen from equation (35) $\tau(t)$ is computed by substituting $\ddot{q}_d - u$ for $\ddot{q}$ in (27), that is, buy solving *inverse dynamic problem*. It is evident that the performance of this control method heavily depends on the accuracy of the system model. Generally speaking, experts agree that there are two main factors that complicate the practical application of the CTC method for controlling manipulators. These are the complexity of the system being controlled and the difficulty of developing an accurate dynamic model of the system. Regarding the challenges related to system complexity, it should be understood that, in most cases, the issue lies in the complexity of the system's mathematical model and the difficulties in working with it, as well as (in some cases) computational challenges that arise during the implementation of controllers. We believe that the use of Lie theory-based kinematic and dynamic models for manipulators can significantly mitigate these problems.

IV. Conclusion

In Lie theory, continuous transformation groups are treated as geometric objects, namely smooth manifolds. Rigid body motions have a Lie group structure, which makes Lie theory based kinematic models of multilink manipulators both natural and simple. Lie theory provides a mathematical framework that allows for formulating kinematics and dynamics in a coordinate-invariant form. In this paper, based on Lie theory, an inverse robot dynamics model is established to simplify the computation of inverse dynamics for serial manipulators. Calculating the inverse robot dynamics is complex, and as is well-known, this is one of the main obstacles in the development of computed torque control. We believe that the application of Lie theory methods can help address this problem.

Acknowledgment

This research has been funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. AP13068317).

References

[1] C. Calderón-Cordova, R. Sarango, D. Castillo, V. Lakshminarayanan, "Deep Reinforcement Learning Framework for Control of Robotic Manipulators in Simulated Environments", Article in IEEE Access, vol. 12, July 2024, pp. 103133-103133.

[2] A. Sáfár, B. Beszédes, "Educational Aspects of a Modular Power Management System", Conference: AIS 2019: 14th International Symposium on Applied Informatics and Related Areas organized in the frame of Hungarian Science Festival 2019 by Óbuda University, 2019.

[3] G. Györök, B. Beszédes, "Duplicated Control Unit Based Embedded Faultmasking Systems", IEEE - Szakál, Anikó; IEEE (szerk.) 2017 IEEE 15th International Symposium on Intelligent Systems and Informatics (SISY): 15th IEEE International Symposium on Intelligent Systems and Informatics. New York, Amerikai Egyesült Államok: IEEE (2017), pp. 283-288.

[4] G. Györök, B. Beszédes, "Fault tolerant power supply systems", Orosz, Gábor Tamás (szerk.) 11th International Symposium on Applied Informatics and Related Areas (AIS 2016). Székesfehérvár, Magyarország: Óbudai Egyetem, 2016, pp. 68-73.

[5] J. H. Kim, "Induced Norm-Based Analysis for Computed Torque Control of Robot Systems", Article in IEEE Access, vol. 4, 2016, pp. 1-11.

[6] T. Yaren,S. Küçük "Dynamic modeling of 3-DOF rrp type serial robotic manipulator using lagrange-euler method", International Marmara Sciences Congress (Autumn) 2019 Proceedings Book (Natural and Applied Sciences), 2019, pp. 372-378.

[7] V. Duruisseaux, T. Duong, M. Leok, N. Atanasov, "Lie Group Forced Variational Integrator Networks for Learning and Control of Robot Systems", Proceedings of Machine Learning Research vol XX:1–21, 2023, pp. 1-21.

[8] A. Muller, Sh. Kumar, T. Kordik "A Recursive Lie-Group Formulation for the Second-Order Time Derivatives of the Inverse Dynamics of Parallel Kinematic Manipulators", IEEE Robotics and Automation letters. preprint version. April, 2023, pp. 1-8.

[9] R. J. Salman, H. M. Alwan, M. A. Yousif, A. M. Abdullah, "Computed Torque-NN-GWO Dynamic Hybrid Control of Manipulator Robotic Arm", Iraqi Journal of Computers, Communications, Control & Systems Engineering (IJCCCE), vol. 24, no. 2, June 2024, pp. 1-16.

[10] A. Charles, "Introduction to Computational Algebraic Geometry in Robotics", Research Invention Journal of Engineering and Physical Sciences, vol. 3, no. 1, 2024, pp.70-79.

[11] N. E. N. A. Zakey, M. H. M. Zaman, M. F. Ibrahim, "Structural Optimization of 4-DOF Agricultural Robot Arm", Jurnal Kejuruteraan, vol. 36, no. 3, 2024, pp. 1127–1134.

[12] H. D. Salman, M. N. Hamzah, S. H. Bakhy, "Kinematics analysis and implementation of three degrees of freedom robotic arm by using MATLAB", The Iraqi Journal For Mechanical And Material Engineering, Vo2l, No.2, June 2021, pp. 118-129.

[13] R. Alizade, "Structural Synthesis of Robot Manipulators by Using Screw with Variable Pitch" Universal Journal of Mechanical Engineering vol. 7, no. 2, 2019, pp. 50-63, 2019, pp. 50-63.

[14] E. J. Haug, A. Peidro, "Redundant Manipulator Kinematics and Dynamics on Differentiable Manifolds", Journal of Computational and Nonlinear Dynamics, vol. 17, 2022, pp. 111008-1-111008-15.

[15] N. Alhousani, M. Saveriano, I. Sevinc, T. Abdulkuddus, H. Kose, F. J. Abu-Dakkae. "Geometric Reinforcement Learning: The Case of Cartesian Space Orientation", Geometric Reinforcement Learning For Robotic Manipulation, 2022, pp. 1-9.

[16] W. Khalil, "Modeling and Control of Manipulators - Part I: Geometric and Kinematic Models", HAL Id: cel-02129939, May 2019.

[17] A. Mueller, "Recursive second-order inverse dynamics for serial manipulators," in 2017 IEEE International Conference on Robotics and Automation (ICRA), 2017, pp. 2483–2489

[18] E. Latif, R. Parasuraman, "On the Intersection of Computational Geometry Algorithms with Mobile Robot Path Planning", Algorithms, vol. 16, no. 498, 2023, pp. 1-15.

[19] W. Zhang, Z. Shi, X. Chai, Y. Ding, "An Efficient Kinematic Calibration Method for Parallel Robots with Compact Multi-DOF Joint Models", Journal of Mechanisms and Robotics, January 2024, pp. 1-33.

[20] K. M. Lynch, F. C. Park, Modern robotics. Cambridge University Press, 2017.

[21] Loring W. Tu, An Introduction to Manifolds, Universitext, 2008.

Presentation of an Educational Mobile Application for Learning the Theory of Cutting Technology

Csaba Bráda
Alba Regia Faculty Obuda University
Székesfehérvár, Hungary
brada.csaba@amk.uni-obuda.hu

Monika Pogatsnik
Alba Regia Faculty Obuda University
Székesfehérvár, Hungary
pogatsnik.monika@amk.uni-obuda.hu

***Abstract*— Our study presents a self-developed phone application, which helps university mechanical engineering students learn the material related to cutting theory. Nowadays, countless applications for calculating the technological parameters of cutting are available on the world wide web, but almost all of them are characterized by the fact that they do not require knowledge of relationships. The application helps to understand these relationships, to see how the individual cutting parameters interact and build on each other. The material also includes possible directions for development.**

Keywords: mobile application, education, cutting

I. Presentation of the purpose and concept of the application

The purpose of the application is to help engineering students understand and master the theoretical relationships of machining [1], and at the same time to provide data that can be used in the design of industrial processes.

Provide an overview of the hierarchy and interdependence of process parameters, as well as provide an actual, numerically correct solution for a cutting machining task.

The implementation of the application required the harmonization of three essential areas:

1. setting up a hierarchical system of technological parameters, interpreting the system of their relationships,
2. applicability, provision of access,
3. actual programming and its framework.

A significant part of the manufacturing processes used in mechanical engineering belong to tertiary forming technologies, especially cutting processes [2]. The planning and implementation of the production of a finished product created by separating the surplus from a larger raw material requires a high degree of professional knowledge. The quality of the machined material, its other properties and the effect of physical characteristics during machining on the material must be known.

In addition to defining numerical data, engineers are expected to know and understand the logic of machining, the system of its relationships, and the mathematical formulas describing these relationships.

These connections are learned by different engineers in different depths – content – during their studies. A production design engineer thinks differently than a CAD-CAM IT engineer when they create the same production process.

The application includes four cutting processes: turning, drilling, face and mantle milling.

Nowadays, the way information is obtained has shifted significantly towards information technology. The advent of computers and the subsequent spread of mobile devices, complemented by the expansion of the availability of the World Wide Web and cloud services, have resulted in the fact that obtaining certain information is not limited either locally or in time [3], [4], [5], [6].

Taking all this into account, it was decided to configure the application for the mobile phone interface.

The application is written in Appinventor [7], [8], [9] development environment, which I recommend for anyone who wants to create their own application without deep programming knowledge.

In programming, design and function are separated. That is, in the first round, the appearance of the application can be designed, to which a given function can be assigned later. Although the system does not require programming knowledge, it does require programming logical thinking, which differs to a greater or lesser extent from so-called logical thinking. Rather, this approach requires a cause-effect mindset.

Options for further development of the application:

- expansion of cutting processes,

- occasional shuffling or expansion of the reply to buttons, and
- show comments

II. New content of the application

After reading the introduction, many people will surely wonder if such an application has already been created and is available to those interested. Indeed, countless applications and internet availability deal with cutting technologies. However, in most cases, these applications are made by cutting tool companies to help with tool selection and are designed for practicing engineers. That is, the goal is not to help learning, but to provide the result of a technological parameter related to a given machining. A typical example is shown in Fig. 1 [10].

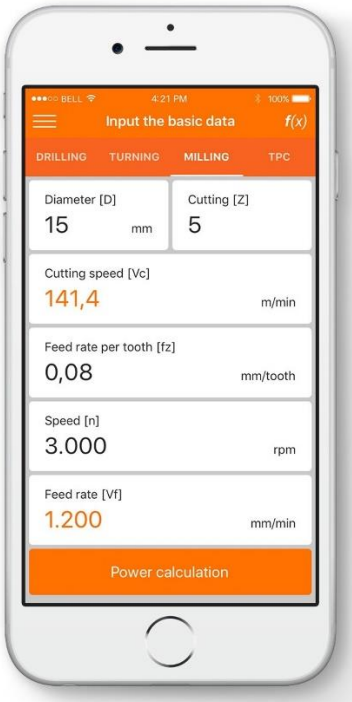


Fig. 1: Technology data calculation app

Calculation formulas are also available in several user interfaces, such as an annex to Fig. 2 [11] or catalogs.

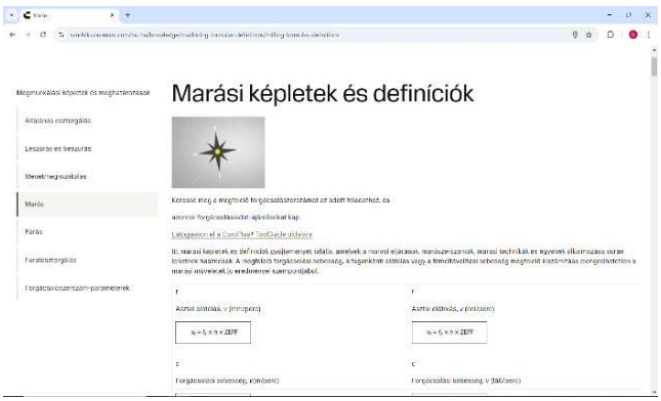


Fig. 2: Cutting formula collection

An additional industrial phone application is available, a tool for engineers and those involved in machining processes. Our application gives users access to a wide range of calculations and functions that simplify the workflow, increase accuracy and increase productivity in many machining operations, including: turning, milling, drilling, tapping, thread milling, center drilling, punching, knurling. and more, pl. MaCa – Machining Calculator – Sandvik Coromant product [13], or Microbor app – Soft Engineer Developer product [14], or CNC Machinist Calculator Ultra [15]

In each application, the primary goal is to determine the result. The program assumes that the user knows the theoretical knowledge and connections, so it does not cover them.

From the point of view of learning, it is unfavorable that

- The necessary information requires the simultaneous use of several interfaces, and
- It does not make the user think, does not put him in a decision-making position and therefore has neither positive nor negative reinforcement.

The novelty of the application to be presented is therefore that:

- integrates the formula collection and the solution and result determined by calculation onto one interface,
- creates a decision using multiple choice methods, in which the user must make a specific statement,
- in case of the right decision, it gives the result in a prominent form as positive reinforcement,
- When determining basic data, it provides textual support – plus information – to the user.

In my opinion, one of the crucial elements of the learning process is repetition. The purpose of repetition is to transfer the perceived information from short-term memory to longterm memory, where it is stored permanently and can be recalled as needed. The use of the application is intended to facilitate this kind of short-term to long-term memory transfer through self-regulated repetition.

III. INTRODUCTION TO THE APP.

After downloading the application or starting the program, the selection page opens, where the user can choose from the options for calculating the parameters of the three most used cutting technologies. The three main categories are:

- turning,
- milling, • drilling.

Among the options for milling technology, two types of machining can be imagined, which show significant differences in the calculation methodology.

To avoid the need to create a selector page on the outside, the two types of milling have separate marking and selection buttons. The image of the starting page is shown in Fig. 3. The header contains help on how to use it.

Fig. 3. Application start page

The selectors consciously show the workpiece-tool relationship typically associated with machining. In this case, the transfer of knowledge is not targeted, it is exact, but it helps to connect theory and practice in the sense that the user can decide whether the machining technology seen in a real situation is face milling or mantle milling.

The choice of technology is followed by the calculation of the parameters of machining. In this tutorial, choose turning as a technology. Clicking on the selector button displays the calculation page. Fig. 4.

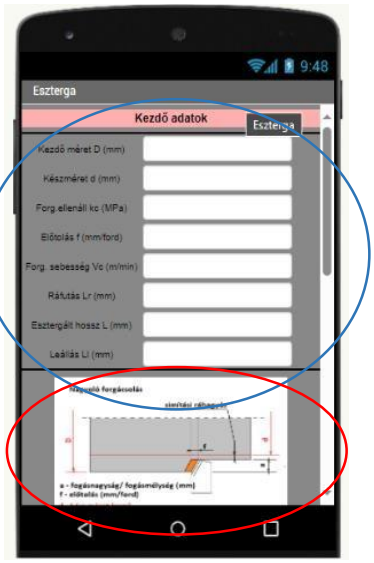


Fig. 4. Initialization

The calculation side is divided into three parts. In the first part (marked with a blue circle) the basic data of machining is entered. Here, the user can enter more—less, arbitrary data. This data can be completely fictitious, but it can also be data about an actual task.

It is important to note that for some data, when you select a text box, a pop-up window will tell you about possible values. For example, the cutting speed for high-speed steel blades should not exceed 40 m/min.

In the second part (area marked in red), the user receives graphic information about the characteristics of the machining, their marking and meaning.

The third part contains the actual calculations, which are shown in Fig. 5.

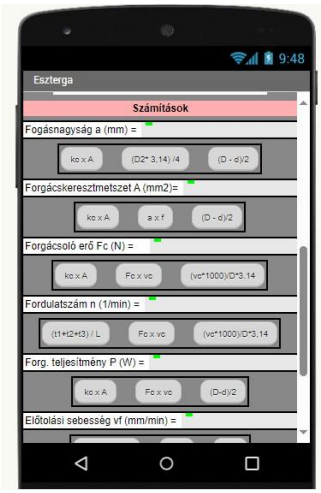


Fig. 5. Calculations

On the calculation page, the name, sign and unit of measurement of the calculated parameter are displayed, as well as three formulas, The student's task is to choose the correct formula to use for calculation from the three formulas.

If the answer is incorrect, nothing happens, but the user cannot proceed.

If the answer is correct, the solution will be written next to the question in a luminous green frame. It would be easy for the wrong answer to create some kind of spectacular effect, but that's not the goal. I find positive reinforcement more useful.

After answering all the questions, you can return to the opening side, where you can start a new program or exit the application.

The logic shown in turning is displayed for each machining method.

IV. THE PROGRAMMING INTERFACE

ELTE lists 228 programming languages [12]. Each of them has its own system, syntax. Without knowledge of these rules, it is impossible to carry out programming. Appinventor or Scribus, on the other hand, does not require possession of this knowledge. You simply need to assign a function to specific elements. The program translates these relationships into a given language, without external intervention within the program.

In the case of appinventor, programming takes place in two steps. The first step is to design the image, practically in a graphical environment. You can determine the position of individual images, buttons, controls, e.g. input fields. See Fig. 6.

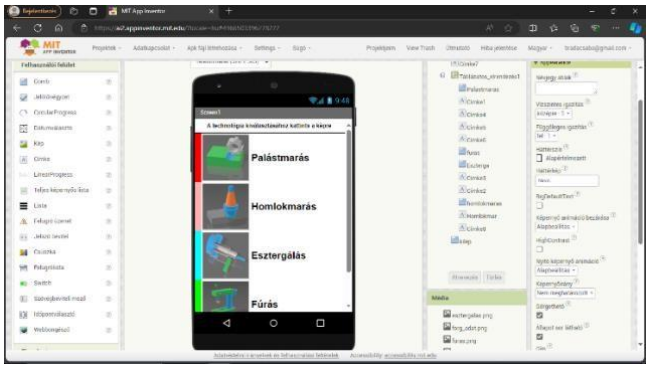


Fig. 6: Graphical representation of a program.

The user interface can be used to place the elements shown in Fig. 7 (highlighted in the left part of Fig. 6), which contain practically the entire range of functional elements. From these elements the entire appearance and formal image of the application can be built.

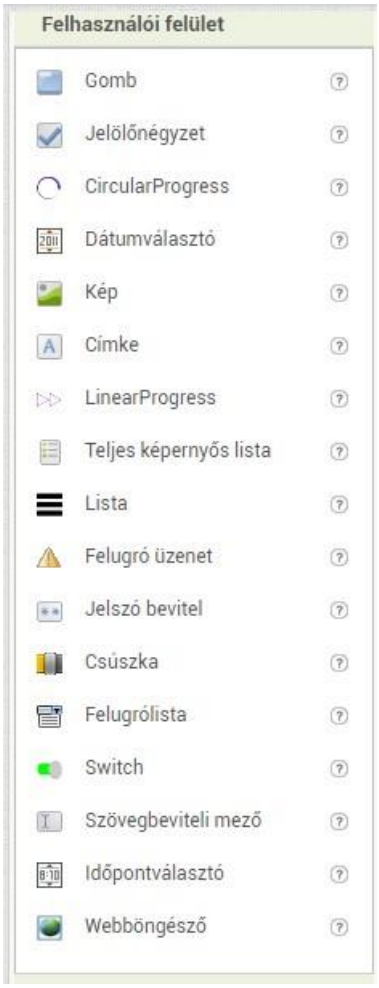
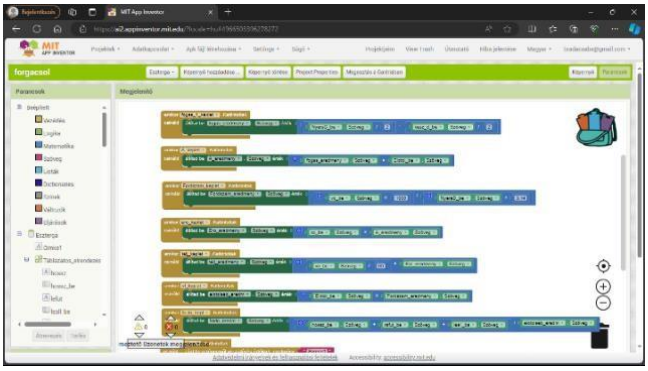


Fig. 7. Widgets

However, appearance is only an image, part of the overall impression, but not yet a workable program. It must also be possible to assign functions to the displayed items, which are accomplished by another element of appinventor, the system of command lines. The given graphic element must be given a task, The task page is shown in Fig. 8.



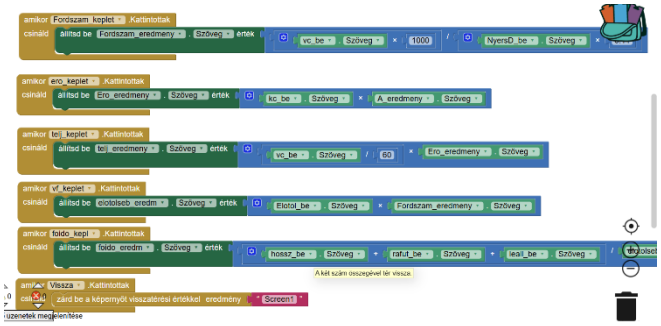


Fig. 8. Function commands

Fig. 9 highlights the types of command categories that can be selected. Clicking on the selected item opens additional commands that can be categorized Fig. 10, which can be placed on the editing page using a simple drag-and-drop method.

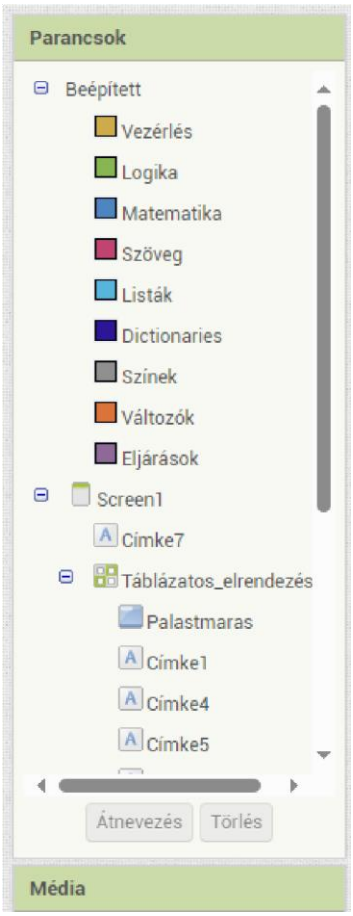


Fig. 9. Built-in elements

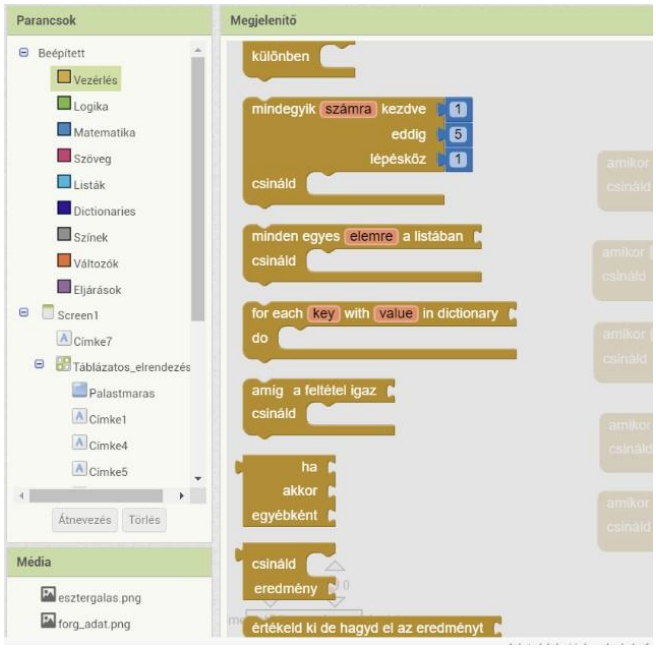


Fig. 10. Collection of optional commands

What is important is that writing a program does not require the components to be in the correct order. The various command blocks are suitable for running independently, regardless of the software. A command block is created by concatenating different commands, which requires only knowledge of the logic of operation from the programmer. Fig. 11 shows part of the program for turning, where each block contains a control command element, mathematical calculation, text input, or display.

Fig. 11. Turning program detail

Within the editing program, it is also possible to test the completed program parts on a virtual phone

V. Conclusion

The completed application will be tested in the first semester of the 2024/25 academic year among second-year mechanical engineering students. However, regardless of the test results, there are already a few observations that can be made:

- changes to the appearance of the application should be made that are more in line with the expectations of the target audience, i.e. those in secondary and higher education. You need a little "life" in appearance.
- Important or interesting information should be made available in several places, such information can be, for example, an article published in a professional journal in pdf format, or an internet link to a demonstration video or tool catalogue.
- It shall be possible for answers to a given question to appear in random order at each startup, thus avoiding routine-like solutions.

About the programming interface, we can only recommend it to anyone who wants to create an application that is quick and easy or does not have sufficient experience in using a traditional programming language.

REFERENCE

[1] Cs. Bráda, Forgácsing Szega Books Ltd. Pécs, 2023.

[2] W. Grzesik, Advanced machining processes of metallic materials: theory, modelling and applications. Elsevier, 2008.

[3] J. Hense, H. Mandl, C. Grasel, "Probléma-központú tanulás" Információs Társadalom, 3(2), 52-62., 2023.

[4] N., Tolner, M. Pogátsnik, "The large language models in the service of education", Magyar Nyelvőr, 148. Special Issue in English (2024-5) pp. 629-649. , 20 p., 2024.

[5] J. Módné Takács, M. Pogátsnik, "The Presence of Cybersecurity Competencies in the Engineering Education of Generation Z", Acta Polytechnica Hungarica 21, 6 pp. 107-127. , 21 p., 2024. [6] J. Módné Takács, M. Pogátsnik, "The online learning from the students' perspective" In: Szakál, Anikó (szerk.) IEEE 19th World Symposium on Applied Machine Intelligence and Informatics (SAMI), Budapest, Magyarország : IEEE Hungary Section (2021) 507 p. pp. 00027-00032. , 6 p., 2021.

[7] Massachusetts Institute of Technology, Appinventor, 2012-2024, https://appinventor.mit.edu/

[8] W. B. Sabado, "Education 4.0: Using Web-based Massachusetts Institute of Technology (MIT) App Inventor 2 in Android Application Development", International Journal of Computing Sciences Research, 8, 2766-2780. 2024.

[9] E. W. Patton, M. Tissenbaum, F. Harunani, „MIT app inventor: Objectives, design, and development", Computational Thinking Education, 31-49. 2019.

[10] Hoffmann Group, Machining kalkulátor, https://www.hoffmanngroup.com/HU/hu/hohu/szolgaltatasok/eszkoz-es-adatkereso/machining-kalkulator/e/37321/

[11] Sandvik Magyarország Kft., Marási képletek és definíciók, https://www.sandvik.coromant.com/hu-hu/knowledge/machiningformulas-definitions/milling-formulas-definitions

[12] ELTE IK, Programozási Nyelvek és Fordítóprogramok Tanszék, http://nyelvek.inf.elte.hu

[13] https://play.google.com/store/apps/details?id=com.sandvik.materialcalculator.activities&hl=hu

[14] https://play.google.com/store/apps/details?id=com.millingcalculator.millingcalculator&hl=hu

[15] https://play.google.com/store/apps/details?id=com.pluckedstudios.cncmce&hl=hu

Simulation and Analysis of Interference in Wireless Data Networks

Aslima Alimkhanova
Faculty of Technology, Department of Information Technology
K.Kulazhanov Kazakhl University of Technjljgy and Business
Astana, Kazakhstan
aslima_alimhanova@mail.ru

Svetlana Grigoryeva
School of Digital Technologies and Artificial Intelligence
D. Serikbayev East Kazakhstan Technical University
Ust-Kamenogorsk, Kazakhstan
sgrigorieva@inbox.ru

Nikita Pronevich
School of Digital Technologies and Artificial Intelligence
D. Serikbayev East Kazakhstan Technical University
Ust-Kamenogorsk, Kazakhstan
sevenyt@mail.ru

Yersin Kozhanzharov
School of Digital Technologies and Artificial Intelligence
D Serikbayev East Kazakhstan Technical University
Ust-Kamenogorsk, Kazakhstan
kozhanzharov_ersin@mail.ru

***Abstract*—The article is devoted to the development of a software and hardware platform for generating and modulating interference in wireless data transmission channels. The research aims to improve understanding of the impact of interference on wireless networks and develop methods to control it. The creation of hardware and software solutions makes it possible to simulate various interference scenarios, analyze and test them, which helps to increase the stability and reliability of wireless communications. The transmitter, receiver and interference generator are developed on the basis of the ESP32 microcontroller. To control the parameters of each device and simulation scenarios, software was developed in C++. A software module is presented for analyzing amplitude modulation and simulating the effect of possible interference on the signal when transmitting data in Wi-Fi networks. The proposed platform architecture allows for the implementation of other types of modulation.**

Keywords – simulation, interference, amplitude modulation, wireless network, Wi-Fi, software.

I. Introduction

Wireless technologies are of great importance in the modern world for ensuring communication between people, devices and systems. The relevance of wireless communications is explained by its wide range of applications, convenience, flexibility, cost-effectiveness and ability to support modern technologies and innovations. With the rapid development of technology and the increase in the number of connected devices the role of wireless communications will only increase [1-3], contributing to further progress and improving the quality of life of people.

The reliability of wireless networks impacts the efficiency and security of systems ranging from consumer IoT to mission-critical industrial and medical applications. With the increase in the number of devices and the volume of transmitted data, the problem of interference in wireless channels becomes relevant [4-5]. Interference can lead to failures in data transmission, deterioration in communication quality, and even complete system failure. The introduction of new anti-interference techniques and modulation technologies opens up opportunities to improve the reliability and performance of wireless networks. In addition, due to the rapid transition to 5G and the creation of new next-generation networks, the study of interference and methods for eliminating it has become strategically important [6 -8]. Scientists and engineers will be able to better understand and manage wireless communications components through the research and development of both hardware for generating and modulating interference and software for analyzing and modeling interference. The authors conducted research related to the analysis of problems in modern wireless networks [9], the impact of interference in optical wireless networks based on simulation modeling [10-11].

This work set the task of implementing a hardware platform for generating, modulating and analyzing interference in wireless data transmission channels, as well as developing software for modeling and analyzing this interference. The goal of this research is to improve our understanding of the processes that occur in wireless networks as a result of interference, and to find ways to effectively control and minimize these processes.

II. Platform for Simulation of Interference in Wireless Data Networks

A. Wireless Operation Principle

The operating principle of wireless communications is based on the transmission of data via electromagnetic waves through airspace. This process can be broken down into several key steps. In the first stage, the data to be transmitted is converted into a form suitable for transmission over a wireless network. This is done through modulation, a process in which information, such as a digital signal, is converted into an analog signal with a specific frequency and amplitude. There are the following types of modulation - amplitude modulation, angular modulation including frequency and phase modulation, and pulse modulation. A device that models, amplifies and transmits information over the air is called a transmitter. The transmitter converts the analog signal into electromagnetic waves, which in turn are received by the antenna. Electromagnetic waves propagate in the air and move freely in space. At this stage, the signal may be subject to various types of attenuation and interference, including absorption, reflection, diffraction and scattering.

One of the important aspects between the transmitter and the antenna is their matching, as well as the choice of data transmission frequencies.

The receiver receives the transmitted signal using an antenna and converts it back into an electrical signal. The antenna picks up electromagnetic waves and transmits them to the receiving device, where the demodulation process occurs. The electrical signal is converted back to the original data. This may be an audio signal, video signal, text information or digital data. After demodulation, the resulting data can be processed and used in various applications. For example, in cellular networks, data can be converted into voice information or text messages, in Wi-Fi networks into Internet traffic, and in Bluetooth connections into commands for controlling devices.

Thus, the operating principle of wireless communication includes modulation, transmission, propagation, reception and demodulation of the signal, providing efficient and reliable data transmission without the need for physical connections. This process allows devices to exchange information over long distances and in environments where using wired technology would be inconvenient or impossible.

To ensure reliable data transmission and minimize the impact of signal interference, it is important to use modern signal control and filtering techniques [12]. The effective use of adaptive filters and noise reduction technologies can significantly improve the quality of communications. Adaptive filters are able to dynamically change their parameters depending on environmental conditions and the level of interference, which makes them ideal for use in mobile communications and radio broadcasting. In addition, the use of various types of coding, such as forward error correction coding, can significantly reduce the impact of errors caused by interference by restoring the original data even if it is partially lost during transmission [13].

Such techniques not only improve data reliability, but also improve overall system performance by reducing the number of retransmissions required. Also important is the use of channel diversification, when several frequencies or channels are used simultaneously to transmit the same information, which significantly reduces the risk of complete data loss due to interference on one of the channels [14-15]. The introduction of intelligent spectrum management systems that are capable of analyzing the surrounding radio-electronic space and optimally distributing available resources is also a promising direction in modern radio engineering [16]. These systems can prevent frequency conflicts and automatically select the least congested channels, reducing the possibility of interference. All these technologies and methods together form a powerful basis for creating highly efficient and reliable communication systems that can withstand a variety of interference and ensure stable communication quality in any conditions.

B. Description of the hardware and software platform

To implement a hardware and software platform for research and analysis of interference in wireless networks, it is necessary to design a transmitter, receiver and interference generator. The transmitter must generate a stable signal and create a Wi-Fi network, which will simulate real-life wireless network conditions. The receiver must accurately register the incoming signal and the signal from the interference generator. The interference generator, as a key element of the design, must generate a variety of types of interference in order to effectively analyze the robustness of various communication protocols to external influences.

In addition to hardware development, the creation of software for device control and data processing is of great importance. The software must be as flexible as possible to allow users to change test parameters such as frequency, signal strength, and interference characteristics. In addition, it is necessary to create a user interface for the platform that is easy to use and allows quick access to test results. Integration and testing of each part of the project are important steps. This includes checking device compatibility and testing performance under different operating conditions and load levels.

The project architecture is presented in Figure 1.

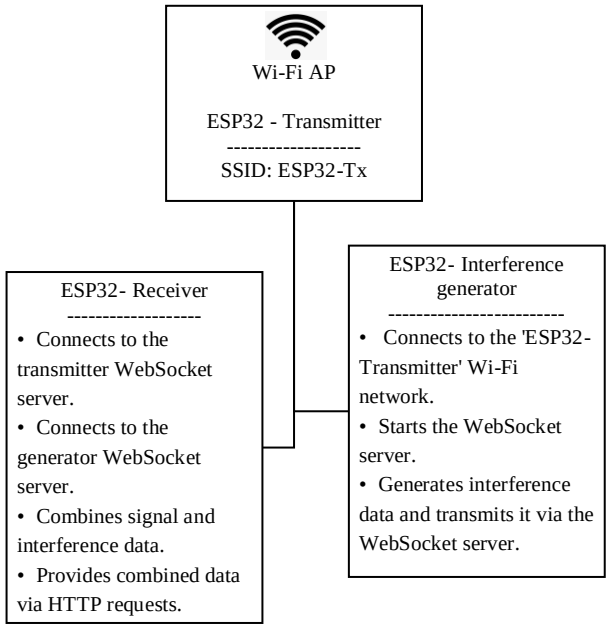


Fig. 1. Project architecture

The transmitter, receiver and jammer are developed based on the ESP32 microcontroller. The devices perform different functions. The transmitter creates a Wi-Fi access point, raises an HTTP server to serve web pages, and a WebSocket server to transmit signal data. It generates sine wave data and transmits it via WebSocket. The interference generator connects to the Wi-Fi network created by the transmitter and raises a WebSocket server to transmit interference data. It generates different types of interference depending on the mode selection and sends data via WebSocket. The receiver also connects to the transmitter's Wi-Fi network, connects to the transmitter's WebSocket server to receive signal data and to the interference generator's WebSocket server to receive interference data.

The interaction between the components occurs as follows:

– the interference generator and receiver are connected to the Wi-Fi network created by the transmitter,

– the transmitter generates and transmits signal data through a WebSocket, and the interference generator generates and transmits interference data through its WebSocket,

– the receiver receives signal data from the transmitter and interference data from the interference generator, combines this data and provides it via HTTP requests.

The proposed architecture allows you to flexibly control interference modes, monitor signal changes and analyze the effects of interference in real time.

For this project, the task was to create a working system with reliable and stable software operation.

The transmitter in this system is responsible for generating sine waves that mimic analog signals in digital form. The main task of the transmitter is to send data regularly. The algorithm of operation of the transmitter is shown in Figure 2.

The developed program initializes the Wi-Fi connection, sets up IP addresses and starts the server. The data to be sent is generated and packaged in accordance with the transmission protocol, after which it is transmitted to the receiver. Special attention was paid to transmission efficiency and minimizing delays.

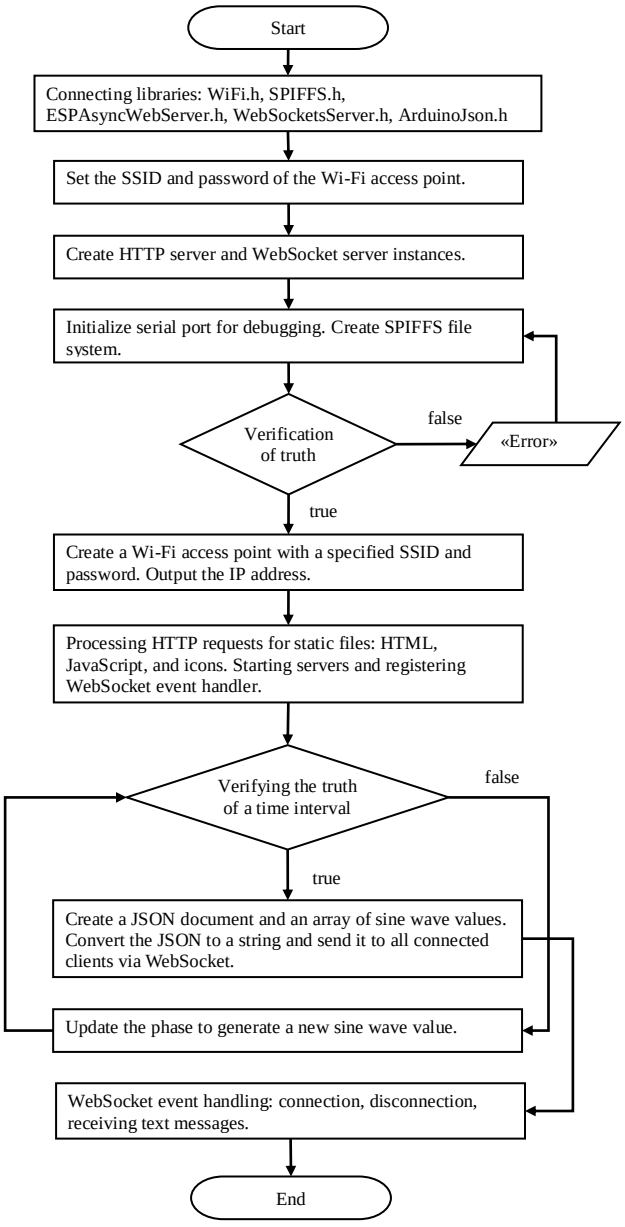


Fig. 2. Transmitter operating algorithm

The receiver is used to receive data from the transmitter and, in some cases, from the interference generator. The program on this device is configured to receive data via Wi-Fi. This data is then processed and visualized as a graph on a web page, allowing the user to observe the signal processing process in real time. The received data is analyzed for errors and losses, which makes it possible to assess the quality of wireless communications. The analysis results are converted for subsequent study and comparison with theoretical interference models. The receiver operating algorithm is shown in Figure 3.

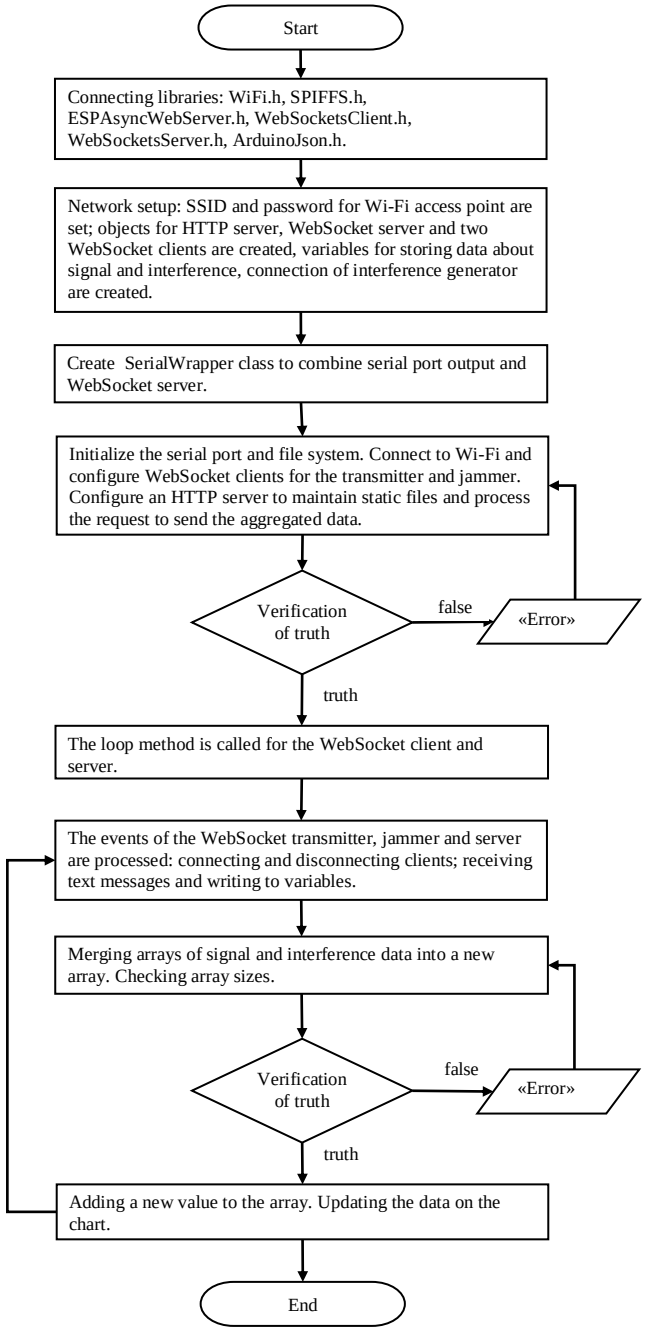


Fig. 3. Receiver operatin galgorithm

The interference generator is the main component of the system. It creates artificial interference by simulating various conditions that may be encountered in the operation of wireless devices. Interference is generated according to a specified algorithm, which can be configured in real time through the user interface.

This allows you to simulate various network interference conditions, including changes in signal strength and channel switching. The interference generator connects to a Wi-Fi network to be able to send data to the receiver. In this mode, it acts as a client on the network. WebSocket is used to send interference data to the receiver. This allows you to transmit data in real time and observe its effect on the processed signal. The code is responsible for generating random numbers that simulate interference and sending this data via WebSocket. The interference is generated as random values in a given range and can be adapted to simulate various types of noise or distortion. The implementation of the noise generator algorithm is shown in Figure 4.

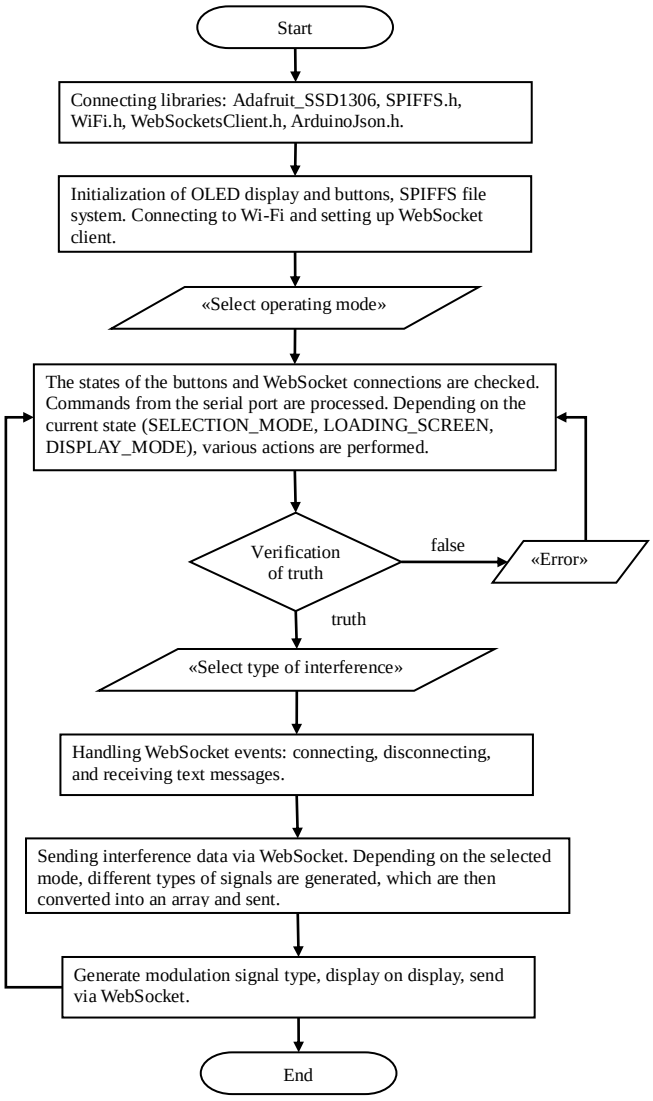


Fig. 4. The algorithm for implementing the interference generator

One of the difficulties was synchronizing the operation of all three devices. It was possible to ensure effective interaction of devices via Wi-Fi using one network. Time synchronization between devices allows you to accurately measure the impact of interference on data transmission, which is critical for the reliability of experiments.

To control and analyze the data transmission process, a web interface has been developed on the transmitter and receiver. This interface is implemented using an HTTP server, which allows users to monitor the transmitted and received data in real time through a web browser.

The transmitter and receiver web interface uses several additional files stored in the SPIFFS file system on ESP32 devices. These files together create an interactive web page that can display and update graphical data in real time, allowing users to monitor the system's operation.

On the transmitter, the web interface provides information about the status of the device and shows the data that is currently being sent. This helps in monitoring transmission activity and allows you to quickly detect any anomalies or changes in data transmission.

On the receiver, the web interface displays information about the received data. This is critical for assessing reception quality and quickly diagnosing interference or data errors that occur during the process.

The introduction of a web interface has improved our system's ability to monitor and analyze wireless transmissions, providing visualization of processes and experimental data.

III. Implementation and Experimental Results

Figure 5 shows the appearance of the developed experimental stand.

Fig. 5. Appearance of the experimental stand

The developed project studies amplitude modulation. Amplitude modulation is one of the oldest and simplest forms of modulation used to transmit information in radio frequency systems. The developed code provides the ability to simulate the most common interference.

The types of interference and their effect on the original signal are shown in Figure 6.

The developed simulator will allow one to observe the effect of interference on the initial signal and evaluate the output signal. This makes it possible to apply the most effective methods and approaches to eliminate interference in wireless data networks.

The main results of the work can be formulated as follows:

1. The developed software allows you to simulate data transmission processes in a wireless network, as well as simulate the impact of interference on this process.

2. The developed hardware platform for generating interference allows you to create different types of interference and vary their parameters to assess their impact on the quality of communication in wireless networks.

3. Flexible architecture and the use of ESP32 microcontrollers have made it possible to create an effective tool for studying the behavior of wireless networks under various interference scenarios.

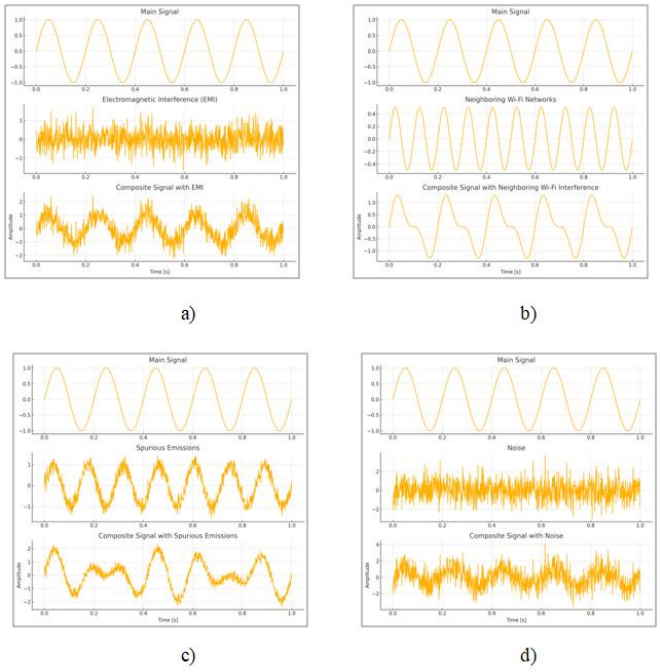


Fig. 6. Simulation graphs of the original signal with various interference: a) electromagnetic interference, b) interference from neighboring Wi-Fi connections, c) spurious radiation interference, d) noise interference

The stages of experimental testing demonstrated the adequacy of the model and the performance of the simulator for analyzing interference in wireless networks.

IV. Conclusion

The proposed platform allows for simulations of the influence of interference on the initial signal in wireless data networks. The hardware platform includes a transmitter, receiver and jammer, which are developed based on the ESP32 microcontroller. Software has been developed to control the parameters of each device and simulation scenarios. The programming language C++ and the development environment Arduino IDE were used. A software module has been implemented to analyze amplitude modulation and simulate the effect of possible interference on the signal during data transmission in Wi-Fi networks. The architecture used makes it possible to implement other types of modulation without the need to change existing connections between program components.

The simulator can be used to conduct research in the field of interference management and optimization of wireless communication systems. Manufacturers of IoT devices, wireless routers, radio modules and other equipment using radio frequency communications can use this platform to test their products, simulating real-life operating conditions. Universities and technical colleges can use the system for educational purposes, allowing students to gain a deeper understanding of wireless communications principles, interference analysis techniques, and interference mitigation techniques. This can also serve as a basis for research and development of new methods for improving communication quality.

References

[1] Global Wireless Connected Devices Market Analysis By Size, By Type, By Application, https://www.linkedin.com/pulse/global-wireless-connected-devices-market-analysis-npybf/.

[2] Number of Internet of Things (IoT) connections worldwide from 2022 to 2023, with forecasts from 2024 to 2033, https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide/.

[3] State of IoT 2024: Number of connected IoT devices growing 13% to 18.8 billion globally, https://iot-analytics.com/number-connected-iot-devices.

[4] Ram Krishan, "Computational Analysis of Interference for Sustainable Operation of Wireless Networks," In book: Sustainable Security Practices Using Blockchain, Quantum and Post-Quantum Technologies for Real Time Applications, 2024, pp. 87-98. doi: 10.1007/978-981-97-0088-2_5

[5] R. Dionísio, F.Ribeiro, J. Metrôlho, "Radio Interference of Wireless Networks and the Impact of AR/VR Applications in Industrial Environments," Electronics, vol.12, issue 1, 2023. https://doi.org/10.3390/electronics12010067

[6] Enoch Adama Jiya, Muhammad B. Muhammad,et all, "Overview on Technologies for Combating Interference and Noise Management in 5G and Beyond Network," Engineering and Technology Journal, 2024, 09(08). doi: 10.47191/etj/v9i08.01.

[7] M.U.A. Siddiqui, F. Qamar, F. Ahmed, Q. N. Nguyen and R. Hassan, "Interference Management in 5G and Beyond Network: Requirements, Challenges and Future Directions," in IEEE Access, vol. 9, pp. 68932-68965, 2021. doi: 10.1109/ACCESS.2021.3073543.

[8] S.H.Ali Kazmi et al., "Routing-Based Interference Mitigation in SDN Enabled Beyond 5G Communication Networks: A Comprehensive Survey," 2023 IEEE Access, Vol.11, pp.4023-4041. doi: 10.1109/ACCESS.2023.3235366.

[9] E. Kozhanzharov, S. Grigoryeva, "Analysis of approaches used to reduce energy consumption in wireless sensor networks," Ust-Kamenogorsk, Kazakhstan, April 2024, part III, pp.150-154.

[10] S.Grigoryeva, A.Baklanov, A.Alimkhanova, A.Dmitriev, Gy.Györök, "Usage of Light-Emitting Diode Lighting and Visible Light Communication Technology for Temperature Control," Acta Polytechnica Hungarica. Budapest. 2021. vol.18, issue 4, pp.7-24. doi:10.12700/APH.18.4.2021.4.1.

[11] S.Grigoryeva, A.Alimkhanova, M. Batalova, "Research of indoor temperature data transmission using visible light communication technology," Journal of Physics: Conference Series, 2021, doi:10.1088/1742-6596/1843/1/012004.

[12] Ge. Gonghao, "Principle and Analysis of Anti-interference and Anti-fading Techniques in Mobile Communication," Highlights in Science, Engineering and Technology, vol.17, pp.336-346, 2022. doi.10.54097/hset.v17i.2625.

[13] M.Tomlinson, C. Tjhai, M. Ambroze, M.Ahmed, M.Jibril, "Error-Correction Coding and Decoding," Signals and Communication Technology, 2017. doi:10.1007/978-3-319-51103-0.

[14] G.Cena, C. Demartini, M.Ghazi Vakili et al., "Evaluating and Modeling IEEE 802.15.4 TSCH Resilience against Wi-Fi Interference in New-Generation Highly-Dependable Wireless Sensor Networks," Ad Hoc Networks, vol. 106, 2020. doi:10.1016/j.adhoc.2020.102199.

[15] I.Levitsky, Y.Okatev, E.Khorov, "Study on Simultaneous Transmission and Reception on Multiple Links in IEEE 802.11be networks," 2020 International Conference Engineering and Telecommunication, 2020.doi:10.1109/ent50437.2020.9431275.

[16] Y.-C.Liang, J.Tan, H.Jia, J.Zhang, L.Zhao, "Realizing Intelligent Spectrum Management for Integrated Satellite and Terrestrial Networks," Communications and Information Networks, vol.6, no.1, pp.32-43, 2021. doi: 10.23919/JCIN.2021.9387703.

Trapped in the Virtual World: The Prevalence of Problematic Internet Use Among Generation Z

Judit Módné Takács
Alba Regia Faculty
Óbuda University
Szekesfehervar, Hungary
modne.t.judit@amk.uni-obuda.hu

Nikoletta Tolner
Alba Regia Faculty
Óbuda University
Szekesfehervar, Hungary
tolner.nikoletta@amk.uni-obuda.hu

Monika Pogátsnik
Alba Regia Faculty
Óbuda University
Szekesfehervar, Hungary
pogatsnik.monika@amk.uni-obuda.hu

***Abstract*— *Background:* A growing challenge for today's young people is problematic internet use. Generation Z, as native users of digital technologies, are particularly vulnerable to the hidden dangers of the digital world. These can range from addiction to lack of privacy.** *Objective:* **This study aims to measure problem Internet use and its underlying factors and investigate relationships between certain dimensions of problem Internet use and demographic factors among Generation Z students in Fejér County.** *Method:* **The study used a questionnaire-based approach. Problematic Internet use was measured on a 5-point Likert scale using a 6-item questionnaire (PIU). A total of N=2438 participants from Fejér County educational institutions (N=38) were surveyed.** *Results:* **Psychometric analyses show the questionnaire has acceptable internal reliability and validity (α=0.642) for interpreting attitudinal dimensions of problematic Internet use in the study sample. The respondents' problematic Internet use is acceptable (M=12.66 SD=3.87), but there are question-specific differences between the groups, especially for girls and the average daily time spent online.** *Conclusion:* **Results indicate that 'neglect' is most highly correlated with average hours spent online. There was also a variable relationship among gender, discipline, average daily internet usage, and total problem internet usage. These findings point to the need for prevention and treatment of problematic use of the Internet in different areas of life.**

Keywords*— *problematic internet use, generation Z, digital addiction, online behavior

I. Introduction

With the rise of digital technology in the 21st century, members of Generation Z have become more vulnerable to the hazards of problematic online usage [1,2]. The COVID-19 pandemic has significantly accelerated the penetration of the online world in the areas of work (home office), leisure (online gaming), and education (online education) [3,4]. More convenient and flexible options have been created for users because of this change. But there are also many negative effects, such as stunted individual development, diminished social skills, online exclusion, and vulnerability to online threats [5]. As young people do not yet have adequate protection mechanisms against the threats of the digital world, it is crucial to develop their digital skills, awareness, and expertise to protect them [6, 7].

Previous studies [8,9,10] have shown that problem Internet use was particularly prevalent in young people, with strong associations with psychological distress, increased stress and depression. Excessive online activity can also affect adolescents' sleep, self-esteem, social relationships, and academic performance [2,11,12]. Mental health problems have serious consequences and are closely associated with severe fatigue, depression, sleep problems, lower quality of life, and can be associated with mental health problems, physical and psychological disorders, and other dependencies, which is why early identification and treatment of problematic Internet use is particularly important.

This study aims to investigate the prevalence and extent of problematic Internet use and related factors among Generation Z students (1995-2009) in Fejér County, from primary to higher education. The research will focus on the duration of Internet use habits, the extent of problematic Internet use and its characteristics. The research aims to provide a comprehensive picture of the dimensions of Internet use among Generation Z students, to provide a more accurate estimate of the extent of problematic Internet use, to identify possible risk factors and to make recommendations for preventing it, or to identify groups more susceptible to becoming addicted to the Internet. To maintain their mental health, understanding problematic internet use is essential. In the region of Fejér County, no comprehensive research on the problematic internet use habits of the Generation Z has been carried out so far, with a comprehensive analysis of the educational institutions.

The following research questions were formulated: (Q1) What is the extent of problematic Internet use in the study population? (Q2) What are the relationships between the dimensions of problematic Internet use and demographic variables?

II. An overview of problematic internet use

Problematic Internet Use (PIU) or Internet Addiction (pathological, maladaptive, excessive, overuse, obsessive Internet use) is a multidisciplinary area of scientific research that is not yet well defined [2]. Problematic Internet use is a behavioral disorder characterized by excessive and compulsive use of the Internet, which has a negative impact on the individual's psychosocial functioning. The clinical relevance of this phenomenon was highlighted by Young [13], who defined it as Internet addiction. Internet addiction [14] refers to a disorder of impulse control that is more severe than that associated with using the Internet to become problematic. Being addicted to the Internet can lead to total isolation, whereas problematic use of the Internet leads to harmful consequences but does not necessarily lead to isolation [15]. In both cases, the user struggles to control how the Internet is used and experiences distress when not using the Internet. Problematic Internet use is a complex phenomenon within which a few types can be identified, including excessive information seeking, gambling addiction, cyber relationship addiction, cybersexual addiction, online shopping addiction, and computer game or VR addiction [2]. Furthermore, one can distinguish between generalized internet addiction, encompassing several dimensions, and specific internet

addiction, focused on one domain such as chat or online games [14].

The FOMO phenomenon associated with the use of social media is particularly prevalent among members of Generation Z. According to the results of a study [16], 40% of respondents regularly experience FOMO, and 5.5% persistently, mainly due to concerns about maintaining social contacts and keeping up to date with current events. According to recent international research, internet addiction is a major social problem, affecting between 7% and 10% of the population, particularly in younger age groups, where it is higher at 19% [17-19]. PIUs are very sensitive to population, methodological, and cultural factors. Generally, PIU is more prevalent among boys, especially regarding online gaming, while girls tend to overuse social media and smartphones [20]. With increasing internet use, the prevalence of PIU is expected to increase in the future [21, 22].

Many psychometric tools are available to measure problematic Internet use (PIU). The IAT (Internet Addiction Test), which is widely used in the diagnosis of PIU, was developed by Young, a pioneer in this field [13]. The CIUS (Compulsive Internet Use Scale) [23] is another widely used measure that focuses on the compulsive nature of PIU. Caplan [24] has developed the GPIUS (Generalized Problematic Internet Use Scale), which examines the psychosocial context of PIU. The common aim of these measures is to quantify the extent of PIU and the characteristics of PIU. In Hungary, the PIUQ (Problematic Internet Use Scale) [25], developed by Demetrovics and colleagues, is mainly used. It consists of 3 subscales (Obsession, Neglect, Control Disorder) and measures different aspects of respondents' Internet use.

III. Research methods

A. Instruments

To measure problematic Internet use, Demetrovics et al [25] developed a 30-item version, which was reduced to 18 items based on psychometric results. In the present study, we use a shortened, validated Hungarian version of this PIUQ questionnaire (PIUQ-SF-6), consisting of 6 items, which measures the participants' level of problematic Internet use on a 5-point Likert scale. The measure assesses the phenomenon in three dimensions (obsession, neglect, control disorder), highlighting key characteristics that are determinants of different forms of addiction. The reliability of the PIUQ-SF-6 is characterized by a Cronbach's alpha of 0.77 [26]. The first dimension measures obsessive thinking (obsession), which refers to the excessive mental presence of Internet activities, especially when the individual feels disabled from performing the desired behavior. The second dimension (neglect) looks at abandonment in other areas of life, including a decline in learning and work efficiency, and a decline in personal and social relationships and fundamental needs such as sleep and eating. The third dimension (control disorder) focuses on compulsivity and disorientation, i.e. the irresistible urge to perform behavior, which manifests itself in an almost compulsive need to perform it. [27]

Responses are used to calculate total scores for each dimension, representing the severity of participants' problematic internet use. The maximum score per dimension is 10, with an overall score of 30. Higher scores indicate more problematic internet use. Respondents are divided into 4 groups, as shown in Table 1, based on the aggregated scores [26].

TABLE I. Interpretation of Problematic Internet Use by group

Group	Method of calculation	Code	Min	Max
no problem group	*average score - one standard deviation below*	NP	0	8.77
average group	*between the mean score and one standard deviation*	AP	8.78	16.53
problem group	*average score + 2*standard deviation between*	PG	16.54	20.41
serious problem group	*average score + 2*standard deviation above*	SP	20.42	30

When assessing the internal consistency of the PIUQ-SF-6 questionnaire used in the survey, the obtained value α=0.642. The obtained Cronbach's alpha value is above 0.6, thus the internal consistency of the questionnaire can be considered acceptable [28,29].

Statistical analysis of the data, normality checks, and various correlation and regression calculations were performed using SPSS software. Analyses are based on anonymous data. Descriptive statistics were calculated, and Spearman's rank correlation was applied. The Kruskal-Wallis H-test was used for the comparison of the medians of several independent groups for the analysis of differences between groups. To compare two groups, the Mann-Whitney U test was used. The statistical significance was determined at the level of $p < 0.05$ [28].

B. Data collection method and sample

An access-based sampling technique was used to collect quantitative data using online and paper questionnaires. The questionnaire was sent to all educational institutions in Fejér County, of which 38 institutions were part of the survey. The research was conducted in late 2023 and early 2024. The response was voluntary and anonymous. The sample was heterogeneous, as the participation of students from primary, secondary, and higher education institutions was variable.

Generation X and Y members outside the target group and incomplete questionnaires were excluded. Finally, N=2438 questionnaires were processed, with 1321 boys (54.2%) and 1117 girls (45.8%) included in the final sample. The proportion of primary school students was 22% (N=532), 65% of secondary school students (N=1573), and 13% of higher education students (N=334). By field of study, 38% of respondents were studying real sciences, 35% human sciences, and 27% general studies. By place of residence, most respondents live in a town or village (N=942, 39% and N=937, 38%), with 21% (N=514) living in the county seat and 2% (N=45) in the capital. The highest educational level of parents is distributed as university or college 41% (N=1005), secondary school leaving certificate 40% (N=980), vocational 16% (N=382), and primary school 3% (N=71).

When examining the proportion of Generation Z boys and girls (N = 2,438), a chi-square test (with Yates correction) using the KSH 2022 Fejér County population age class sex distribution data [30] showed no significant difference (χ^2 = 1.36, df = 1, p = 0.067). Based on our results, the null hypothesis that the gender distribution of the sample does not differ from that of the population is accepted, and the sample can be considered as gender representative for the Z generation in Fejér county.

IV. Results

A. Average time spent online by Gen Z

According to the survey, the average time spent online by Generation Z is 5.05 hours per day (M = 5.047, SD = 2.163). The internet is an essential part of Generation Z's lives, and they spend a significant amount of time online. The results of the Mann-Whitney U-test show that females (mean rank = 1257.29) spend significantly more time online (U = 695566.000, Z = -2.537, p = 0.011) than males (mean rank = 1187.55). Contrary to the common belief that males spend more time online with their digital activities [16], the result in the study population is opposite. In addition, Spearman's rank correlation analysis revealed a positive, weakly significant relationship between average daily time spent online and type of school (r_s=0.123, p<0.01) and field of studies (r_s=0.125, p<0.01).

The results of the Kruskal-Wallis H-test confirm that there is a significant difference between the average time spent on the Internet and the type of school, $\chi^2(2, N=2438) = 56.191$, $p < 0.001$, and between the average time spent on the Internet and the field of study, $\chi^2(2, N=2438) = 60.801$, $p < 0.001$. Students in higher education spend the most time online, on average 5.27 ± 2.02 hours per day (mean rank = 1287.44), almost as much as secondary school students, on average 5.21 ± 2.12 hours per day (mean rank = 1270.66), and the least, 4.21 ± 2.25 hours, as primary school students (mean rank = 1025.22). This difference is likely to be due to differences in learning methods, school and parental screen time, nature of tasks, and social expectations between different types of schools. In terms of subject areas, students studying human sciences spend the most time online with an average of 5.28 ± 2.08 hours (mean rank = 1291.32), students studying real sciences spend almost as much time online with an average of 5.23 ± 2.13 hours (mean rank = 1278.21), and students studying general studies spend the least time online with an average of 4.29 ± 2.22 hours (mean rank = 1044.62). Overall, the different types of education and fields of study are not evenly distributed in terms of Internet use. On average, students in higher education and human and real sciences spend more time online, suggesting that school type and field of study influence individuals' digital media consumption habits.

B. Prevalence of problematic Internet use and its dimensions

The average results for each dimension and group of questions are presented below. The mean score for the first item *("How often do you feel that you should reduce the time you spend on the Internet?")* is 2.70 ± 1.13 (M=2.70, SD=1.13), indicating that participants often feel the need to reduce their Internet use. The second item *("How often do you try to hide how much time you spend on the internet?")* has a mean value of 1.53 ± 0.96 (M=1.53, SD=0.96), meaning that the respondents are less likely to say that they hide or should keep secret that they spend time on the internet. The mean value for the control disorder dimension is 4.22 (M=4.22, SD=1.62) which indicates an average loss of control over internet use for the whole sample.

The dimensions of obsession examine the addiction and emotional dependence associated with Internet use. The first item *("How often do you feel anxious or stressed when you don't use the Internet as much as you would like?")* has a mean score of 1.90 (M=1.90, SD=1.07), while the second item *("How often do you feel depressed, sad, anxious when you don't use the Internet and how often does this feeling go away when you start using the Internet again?")* has a mean value of 1.48 (M=1.48, SD=0.94). The obsession dimension has a mean value of 3.37 (M=3.37, SD=1.76), which is the lowest mean value of the three dimensions.

The dimensions of neglect assess the abandoned behavior caused by internet use. The two items *("How often do you use the Internet when you should be sleeping?"* and *"How often do people around you complain that you use the Internet too much?")* have mean values of 2.98 (M=2.98, SD=1.21) and 2.09 (M=2.09, SD=1.14) respectively. The neglect dimension has a mean value of 5.07 (M=5.07, SD=1.82), indicating that participants often experience neglect symptoms due to excessive internet use, the highest value among the dimensions and the highest variance.

The overall mean score of the PIUQ-SF-6, including all dimensions, is 12.66 (M=12.66, SD=3.88). This indicates moderate problematic Internet use among the population studied. Relatively high variance (15.011) and standard deviation (SD = 3.88) indicate significant individual differences in problem Internet use. Moderate levels of problematic Internet use on the control and neglect dimensions are characteristic of the study sample. The obsession dimensions show less severe problems, suggesting that although Internet use affects participants' lives, levels of emotional obsession and dependence are relatively low.

When examining the correlates of demographic variables of problematic internet use, individuals were grouped into four categories: NP (non-problematic), AP (average), PG (problematic) and SP (severe problematic), which resulted in 12% of participants (N=295) being classified in the non-problematic group, 72% in the average problematic group (N=1765), 12% in the problematic group (N=286) and 4% (N=92) in the severe problematic group. Thus 16% of the study population can be characterized as having problematic internet use.

C. Associations of the PIUQ-SF-6 with demographic factors

Spearman's rank correlations show several significant correlations of different strengths and directions between different demographic indicators, including gender (r_s = 0.150, $p < 0.01$), field of study (r_s = -0.115, $p < 0.01$), average daily time spent online (r_s = 0.287, $p < 0.01$) and Generation Z PIU total score, indicating these variables may influence levels of problematic online use.

The results of the chi-square test confirm a significant association between problematic internet use and gender (χ^2 = 53.430, df = 3, $p < 0.001$). The results for the groups show a higher proportion of males in the NP (non-problematic) and AP (average) categories (NP: male 15%, female 9%, AP: male 75%, female 70%), while females are in the majority in the PG (problematic) and SP (severe problematic) categories (PG: male 8% female 16%, SP: male 3% female 5%). Thus, a higher proportion of female participants are classified in the problematic and severe problematic categories than males.

When looking at the individual dimensions, the results of the Mann-Whitney U-test show that there is a significant difference between males and females on the dimensions of control disorder (U = 613696.500, Z = -7.310, $p < 0.001$), obsession (U = 664122.500, Z = -4.467, $p < 0.001$), neglect

($U = 657593.500$, $Z = -4.694$, $p < 0.001$). Females tend to score higher than males on the dimensions of control disorder (mean rank: female 1330.59 male 1125.57), obsession (mean rank: female 1285.44 male 1163.74), and neglect (mean rank: female 1291.29 male 1158.80), as well as on the total PIH-K6 score (mean rank: female 1333.93 male 1122.74). To sum up, women's problematic use of the Internet has a greater impact on each of the dimensions, especially in the areas of loss of control and neglect of other areas of life. Based on the results of the linear regression analysis ($\beta = 1.145$, $t = 7.347$, $p < 0.001$), women have on average a higher problematic internet use score than men by 1.145 points. Although the model explains a relatively small amount of variance ($R^2 = 0.022$), the result is significant.

Small significant ($\chi^2 = 15.675$, $df = 6$, $p = 0.016$) differences in the distribution of different levels of PIU were observed between studies. The proportion of those in the non-problematic category is slightly higher in the real sciences (real sciences = 14.5%, human sciences = 10.9%, general sciences = 10.3%), while the proportion of those in the problematic category is slightly higher in the human sciences and general sciences (real sciences = 9.2%, human sciences = 12.9%, general sciences = 13.8%).

Looking at the individual dimensions, in the dimension of control disorder ($\chi^2 = 17.978$, $df = 2$, $p < 0.001$), general science students score significantly higher (mean rank = 1288.00) than real science students (mean rank = 1146.42). In the dimension of obsession, there is also a significant difference ($\chi^2 = 22.376$, $df = 2$, $p < 0.001$) in the scores between the different disciplines. In this case, students in the general studies had the highest scores (mean rank = 1289.60), while students in the real sciences had the lowest scores (mean rank = 1139.15). There is also a significant difference in the neglect dimension ($\chi^2 = 10.141$, $df = 2$, $p = 0.006$), where general studies students also scored the highest (average rank = 1283.87), with the lowest being real sciences students (average rank = 1171.13). The results of linear regression show that the subject area has a significant effect on the overall score for problematic internet use ($F = 20.588$, $p < 0.001$), but the effect size is small ($R^2 = 0.008$). Specialization has a negative effect on problematic Internet use level ($\beta = -0.444$, $t = -4.537$, $p < 0.001$), so the level of problematic Internet use decreases slightly with changes in specialization, with slightly lower levels of problematic Internet use among students in the real sciences, and slightly higher levels among students in the general and human sciences. The results suggest that the degree is a weak predictor, but statistically significant.

The results show a significant ($\chi^2 = 291.366$, $df = 12$, $p<0.001$) correlation between the different groups of problematic internet use (NP, AP, PG, SP) and the average daily internet use time categories. The non-problematic (NP) group showed the highest proportion of 2-3 hours of Internet use (36.9%), while the average problematic (AP) group was dominated by 4-5 hours of use (39.9%). The problematic (PG) group showed significant use of 4-5 hours (34.3%) and more than 7 hours (33.6%), while the severe problematic (SP) group showed the highest proportion (40.2%) in the more than 7 hours category. The correlation coefficients ($r = 0.121$, $p < 0.001$, $r_s = 0.055$, $p = 0.006$) indicate a weak but significant positive relationship between average internet use time and PIH groups. Longer Internet use time is associated with a higher probability of belonging to the problematic/serious Internet use categories.

The results of the correlation analysis also show that the more time someone spends online, the more likely they are to manifest problematic Internet use ($r_s = 0.287$, $p < 0.01$), suggesting a significant, moderately strong positive relationship between average daily time spent online and the total number of problematic Internet use scores. A stronger, also significant, relationship was observed for the neglect dimension of problematic internet use ($r_s = 0.359$, $p < 0.01$), meaning that neglecting previously important activities (e.g. studying, working, socializing) is more strongly associated with more time spent online. They are significantly related to the obsession dimension ($r_s = 0.201$, $p < 0.01$), i.e. intensive internet use increases feelings of obsession, dependency, and constant urge to continue. In contrast, no significant relationship was found with the control problems dimension, suggesting that in this dimension the amount of Internet use does not significantly affect problematic Internet use, that the development of compulsivity is not directly related to the amount of Internet use, and that the development of control problems, therefore, depends more on qualitative than quantitative factors.

Average daily time spent online was a significant ($F = 228.90$, $p < 0.001$) predictor of problematic Internet use in the linear regression analysis. The correlation coefficient ($R = 0.293$) indicates a moderate relationship, with $R^2 = 0.086$ indicating that time spent using the Internet explains 8.6% of the variance in PIU. The regression coefficient value ($\beta = 0.525$, $t = 15.129$, $p < 0.001$) is significant and positive, indicating that each additional hour of Internet use increases total PIU by 0.525 points on average. To summarize, the results suggest that increasing daily internet use has a significant, moderate effect on the level of problematic internet use among Gen Z.

V. Discussion

The rapid growth of web usage is affecting all levels of society across the world. Global data show that online presence is becoming an increasingly important part of everyday life, not only for young people but also for older generations. Problematic internet addiction can lead to escaping into the online world, increasing loneliness and depression [15]. On average, people spend 6 hours and 40 minutes online per day, according to a global study [31]. This time varies from a maximum of 9 hours 24 minutes in South Africa to a minimum of 3 hours 56 minutes in Japan. Hungarian users are online for 6 hours 16 minutes a day. The average time spent online per day in the study population is slightly lower than the Hungarian average, with Generation Z spending an average of 5.05 hours online, with differences by gender, school type, and field. Females spend more time online than males, and students in higher education and human sciences spend more time online.

Research on PIU is receiving an increasing amount of attention in the literature. Previous national studies have shown an increasing trend, with 4.3% between 2000 and 2002 [25], a 2013 study showing 7.6% among secondary school students [21], and a 2017 study showing 14% among primary school students aged 10-15 [11]. The focus of research is often on gender, and research has shown that males have higher levels of overall PIU than females [8,11,32,33,34]. Levels of internet use (14%) are associated with gender, with boys more

likely to use the internet excessively, according to a 2017 study of primary school children [11]. A 2021 national survey [35] found that 5.9-7.4% of the Hungarian population aged 18-64 reported problematic internet use, with no significant difference between genders. One-third of 3–11-year-olds and, on average, one in five 12-18-year-olds reported problematic internet use in a survey conducted in 2023 [15]. Females in Generation Z are more likely than males to fall into the categories of problematic internet use and serious problems, according to the present research findings. Overall, the prevalence of problem Internet use was found to be mostly average, but with an important proportion of serious cases (16 %). Thus, there is an increase within Generation Z and a difference in the results of the gender analysis compared to previous research findings. Analyzing the different dimensions, neglect, and loss of control were most affected. Obsession is less common, which is an indication of a lower level of dependency symptoms.

Demetrovics et al [25] showed in their study that problematic Internet use is influenced by several factors. Males tend to neglect more, while females tend to control more. Females scored significantly higher on these dimensions than males. This suggests that, particularly in the areas of behavioral control difficulties and neglect of other activities, women in the study population have higher levels of problematic internet use than men. Individual differences, ranging from mild to severe negative consequences of Internet use, are relatively large.

Previous research has shown that, by age, the youngest age group (9-18 years) scores higher than older age groups on control problems and obsessions [25]. The type of school (primary, secondary, higher education) also influences the development of PIU, based on the association between demographic variables and PIU in the study sample. Furthermore, students with a real science major have a lower risk of PIU, whereas students with a general science or human science major have a higher risk of PIU. Consistent with previous research [8], the more time young people spend online (more than 5 hours per day), the more likely they are to develop problematic Internet use, especially neglect. Previous results according to place of residence have shown that those who live in the capital city have lower scores and those who live in rural areas have higher scores [25]. In the present study, this difference does not appear in Fejér County, where no significant correlation was found between where people live and the overall score for problematic Internet use.

Due to the sampling procedure, the results cannot be generalized to Generation Z at the national level, but for the Fejér County school population, the results are representative of Generation Z. The data are based on self-reporting based on the method of data collection, which may lead to some biases, for example, the willingness of participants to admit real internet use or problems may be limited.

VI. Conclusion and Summary

In the 21st century, the Internet has become indispensable to the lives of young adults, but its overuse can have negative consequences. This study aimed to investigate the problematic internet use habits of members of Generation Z in Fejér County. The study involved 2438 participants in primary, secondary, and higher education. The validated PIUQ-SF-6 questionnaire was used to measure problematic internet use.

The findings show that the average daily time spent on the internet by members of Generation Z is 5.05 hours. On average, university students (5.27±2.02), human sciences (5.28±2.08), and real sciences students (5.23±2.13) spent more time online, indicating that type of school and field of study influence people's digital media consumption habits. The overall mean PIUQ-SF-6 score is 12.66. The results suggest that participants do not need to hide their online presence, although they often feel the need to reduce their internet use. The mean for the dimension of control disorders is 4.22. The highest of the dimensions and the highest variance (M=5.07, SD=1.82), students often experience symptoms of neglect due to excessive internet use. The dimension of obsession, with a mean of 3.37, is the lowest among the dimensions, measuring obsession and emotional dependency related to internet use. Several significant correlations of varying magnitude and direction between the PIU total scores of Generation Z are revealed by the results of the correlation tests. Females have a higher prevalence of problematic use of the Internet, especially in the dimensions of control disorders and neglect. On average, females have 1,145 points more problematic internet use than males. Internet use patterns and problematic Internet use are also influenced by type of school and field of study. Students in higher education spend on average more time on the Internet and are more likely to experience problematic use of the Internet. Students in higher education had slightly lower rates of problematic Internet usage. There was a significant association between the average daily time spent online and problematic Internet use.

The study provides a detailed picture of the different dimensions of problematic Internet use and their correlation with demographic variables and is one of the first to examine problematic Internet use among Generation Z in Fejér County. Young people's mental health, academic performance, and social relationships can be negatively affected by excessive internet use. Problematic Internet use is a complex phenomenon that needs further research. Future research should focus on the young, alpha-gen, the effects of different social media platforms, and mental health issues related to problem internet use. Alternatively, qualitative methods could be complemented by a more in-depth understanding of the experiences and habits of young people.

Acknowledgment

The authors thankfully acknowledge the financial support of this work of Alba Regia Muszaki Felsooktatasert Alapitvany. The research has been supported by Kyndryl Foundation to advance cybersecurity skills and resilience - Growing an Inclusive Cybersecurity Workforce project. The authors declare no conflict of interest. The authors would like to express their gratitude to those who have provided their help in the various experiments.

References

[1] E. Bereczki, “The mysterious Generation Z. Working with today's teens and twenties. - A rejtélyes Z generáció. Együttműködés a mai tizen- és huszonévesekkel”, HVG Könyvek, 2022.

[2] Á. Szapáry et al., “Internet addiction: a medical challenge for the 21st century? - Internetfüggőség: a 21. század orvosi kihívása?”, Orvosi Hetilap, 163(38), 2022, pp 1506-1513.

[3] I. Holik et al., “Under the Spell of Video Games - Minecraft is an Opportunity for Game-based Learning,” Acta Polytechnica Hungarica, vol. 21, no. 3, 2024, pp 305–324, doi: 10.12700/aph.21.3.2024.3.19.

[4] T. Kersanszki et al., “Minecraft in STEAM education – applying game-based learning to renewable energy.,” Deleted Journal, no. 60, 2024, pp 194–213, doi: 10.55612/s-5002-060-008.

[5] Gy. Györök, B. Beszédes, „Artificial Education Process Environment for Embedded Systems.“, 9th International Symposium on Applied Informatics and Related Areas - AIS2014. Székesfehérvár, Magyarország : Óbudai Egyetem, 2014, pp 37-42.

[6] B. Beszédes, "Artificial Education Process Environment for Electro-Pneumatic System," 2023 IEEE 6th International Conference and Workshop Óbuda on Electrical and Power Engineering (CANDO-EPE), 2023, pp 175-180, doi: 10.1109/CANDO-EPE60507.2023.10418034.

[7] Gy. Györök, B. Beszédes, „Adaptive Optocoupler Degradation Compensation in Isolated Feedback Loops“, IEEE 12th International Symposium on Applied Computational Intelligence and Informatics (SACI 2018), 2018, pp 167-172, doi: 10.1109/saci.2018.8440950.

[8] M. Pohl et al., “The Association of Internet Addiction with Burnout, Depression, Insomnia, and Quality of Life among Hungarian High School Teachers.”, International Journal of Environmental Research and Public Health, 19(1):438, 2022, https://doi.org/10.3390/ijerph19010438

[9] D. K. Prievara, B. Pikó, “Aimlessly on the World Wide Web - exploring the underlying factors of problematic internet use. - Céltalanul a világhálón-a problémás internethasználat háttértényezőinek vizsgálata.”, Psychiatria Hungarica, 31(2), 2016, pp. 146-156.

[10] R. Velezmoro et al., “Perceived stress, sensation seeking, and college students’ abuse of the internet.”, Comput Hum Behav, 26(6), 2010, pp 1526–1530.

[11] A. Lukács et al., “A study of problematic internet use in primary school pupils aged 10-15. - Problémás internethasználat vizsgálata 10-15 éves általános iskolás tanulóknál.”, Egészségtudományi közlemények: A Miskolci Egyetem közleménye, 7(2), 2017, pp 21-27.

[12] Á. Szapáry et al. “Internet addiction, sleep disturbance, depression, burnout and quality of life among high-school teachers: a path-analytical model. - Internetfüggőség, alvászavar, depresszió, kiégés és életminőség vizsgálata középiskolai tanárok körében: összefüggést elemző modell”, IME: Informatika és Menedzsment az egészségügyben, 21(4), 2022, pp 49-56.

[13] K. S. Young, “Internet Addiction: The Emergence of a New Clinical Disorder.”, CyberPsychology & Behavior, 1(3), 1998, pp 237–244.

[14] A. Gyivicsán, “Generation Z's internet use habits and internet addiction. - A Z generáció internethasználati szokásai és az internetfüggőség.”, Módszertani Közlemények, 63(1), 2023, pp 121–137, https://doi.org/10.14232/modszertani.2023.1.121-137

[15] D. K. Szabó-Prievara, K. Tarkó, “Problematic Internet use among children: - the Internet Addiction Test for Families (IAT-F) in Hungarian. - A problémás mértékű internethasználat gyermekek körében: – az Internet Addiction Test for Families (IAT-F) magyar nyelvű alkalmazása.”, Iskolakultúra, 33(1-2), 2023, pp 77–92, https://doi.org/10.14232/iskkult.2023.1-2.77

[16] J. G. Andres-Pasztor, G. Bak, " Generation Z online - Relationships between social media use, FoMO and social networking - Z generáció online - Közösségi média használat, FoMO és társas kapcsolatok közötti összefüggések" XXIII. Tavaszi Szél Konferencia, Budapest, 2020, doi: 10.6084/m9.figshare.12751049.v1.

[17] G. Tóth, “Internet addiction and its consequences among health care workers. PhD thesis. - Internetfüggőség és következményei egészségügyi dolgozók körében. PhD-értekezés.”, Pécsi Tudományegyetem, Általános Orvostudományi Kar, Klinikai Idegtudományok Doktori Iskola, Pécs, 2021.

[18] D. X. Chia et al., “Prevalence of internet addiction and gaming disorders in Southeast Asia: a meta-analysis.”, Int J Environ Res Public Health, 17: 2582, 2020.

[19] Y. Balhara et al., “Extent and pattern of problematic internet use among school students from Delhi: Findings from the cyber awareness programme.”, Asian Journal of Psychiatry, 34, 2018, pp 38–42.

[20] M. Gao et al., “Internet addiction among teenagers in a Chinese population: Prevalence, risk factors, and its relationship with obsessive-compulsive symptoms.”, Journal of Psychiatric Research, 153, 2022, pp 134–140.

[21] A. Galán, “Development and prevalence of internet addiction. - Az internetfüggőség kialakulása és prevalenciája.”, Metszetek, vol 1, 2014, pp 316–327.

[22] K. Kapus et al., “Prevalence and risk factors of internet addiction among Hungarian high school students.”, Int J Environ Res Public Health, 18: 6989, 2021.

[23] G.-J. Meerkerk et al., "The compulsive Internet use scale (CIUS): Some psychometric properties", CyberPsychology & Behavior, vol. 12, no. 1, 2009, pp 1–6, , doi: 10.1089/cpb.2008.0181.

[24] Caplan, S. E., "Problematic Internet Use and Psychosocial Well-Being: Development of a Theory-Based Cognitive-Behavioral Measurement Instrument," Computers in Human Behavior, vol. 18, 2002, pp. 553-575, doi: 10.1016/S0747-5632(02)00004-3.

[25] Zs. Demetrovics et al., “Presentation of the Problematic Internet Use Questionnaire - A Problémás Internethasználat Kérdőív bemutatása” Psychiat Hung, vol. 19, no. 2, 2004, pp. 141–160.

[26] Zs. Demetrovics et al., “The three-factor model of internet addiction: The development of the Problematic Internet Use Questionnaire.”, Behavior Research Methods, 40, 2008, pp 563–574.

[27] S. Rózsa et al., “Questionnaires, interviews and symptom rating scales used in child psychiatry - A gyermekpszichiátriában alkalmazott kérdőívek, interjúk és tünetbecslő skálák”, EFOP-2.2.0-16-2016-00002 Gyermek és ifjúságpszichiátriai, addiktológiai és mentálhigiénés ellátórendszer infrastrukturális feltételeinek fejlesztése, Budapest, 2020, ISBN: 978-963-9661-55-4

[28] L. Sajtos, A. Mitev, " SPSS Research and Data Analysis Manual - SPSS kutatási és adatelemzési kézikönyv", Aliena Kiadó, 2007. ISBN: 978 963 9659 08 7

[29] L.J. Cronbach, "Coefficient alpha and the internal structure of tests.", Psychometrika 16, 1950, pp. 297–334. https://doi.org/10.1007/BF02310555

[30] KSH, “KSH 2022 Census database: sex, age of population by district - KSH 2022 Népszámlálási adatbázis: A népesség neme, életkora járásonként” KSH (Központi Statisztikai Hivatal), Budapest, 2023. [Online]. Available: https://tinyurl.com/3fkustre (last reviewed: 17/08/2024)

[31] Digital 2024: Global Overview Report, Simon Kemp., 2024. [online] URL: https://datareportal.com/reports/digital-2024-global-overview-report (last reviewed: 17/09/2024)

[32] B. Mustafa et al., “A review of recent research in problematic internet use: gender and cultural differences,”, Current Opinion in Psychology, vol. 36, 2020, pp 124-129, https://doi.org/10.1016/j.copsyc.2020.05.008.

[33] H. Kiss, B. Pikó, “The association of problematic internet use with loneliness among high school and university students. - A problémás internethasználat összefüggése a magányossággal középiskolás és egyetemi hallgatók körében.”, Iskolakultúra: Pedagógusok Szakmai-Tudományos Folyóirata, 27(1-12), 2017, pp 77-85.

[34] H. Odacı, M. Kalkan, “Problematic Internet use, loneliness and dating anxiety among young adult university students.”, Computers & Education, 55. 3, 2010, pp 1091–1097, DOI: 10.1016/j.compedu.2010.05.006

[35] B. Paksi, Zs. Demetrovics, “Addiction problems in Hungary: the situation in the light of population surveys. - Addiktológiai problémák Magyarországon: Helyzetkép a lakossági kutatások tükrében.”, II. Viselkedési addikciók, 2021.

POSTER SESSION II.

Geodesy and Geoinformatics

AI-supported Visual Lisp Programming in Geoinformatics

János Katona
Óbuda University Alba Regia Technical Faculty
Székesfehérvár, Hungary
dr.katona.janos@uni-obuda.hu

Andrea Pődör
Óbuda University Alba Regia Technical Faculty
Székesfehérvár, Hungary
andrea.podor@uni-obuda.hu

***Abstract*— This article describes the development of visual lisp algorithms that have been created with AI support. Geoinformatics software already contains many built-in functions that make the task of users easier. However, by using them together, complex tasks can be performed. The article illustrates the advantages and disadvantages of AI-supported programming through some examples, as well as the steps for refinement. The algorithms are selected from complex editing, computation and subtitling.**

Keywords*— *AI, programming, geoinformatics, visual lisp

I. INTRODUCTION

Artificial Intelligence (AI) has undergone significant transformation since its inception in the mid-20th century. The journey of AI-based programming can be divided into several key phases, each marked by groundbreaking advancements and increasing sophistication.

Early Beginnings and Symbolic AI (1950s-1980s). The concept of AI began with the idea that machines could simulate human intelligence. Early AI research focused on symbolic AI, where programs used symbols to represent knowledge and logic to manipulate these symbols. Notable achievements during this period include the development of the first AI programs like the Logic Theorist and the General Problem Solver, which could solve mathematical problems and puzzles 1.

The Rise of Machine Learning (1990s-2000s). The limitations of symbolic AI led to the rise of machine learning, a paradigm shift where systems learn from data rather than relying on explicit programming. This era saw the development of algorithms such as decision trees, support vector machines, and neural networks. Machine learning enabled significant progress in areas like speech recognition, image classification, and natural language processing 2.

Deep Learning and Big Data (2010s-Present). The advent of big data and advancements in computational power catalyzed the deep learning revolution. Deep learning, a subset of machine learning, involves neural networks with many layers (hence "deep") that can learn complex patterns in large datasets. This approach has led to remarkable achievements, including the development of AI systems that surpass human performance in tasks like image recognition (e.g., ImageNet competition) and game playing (e.g., AlphaGo)3.

Large Language Models and AI Boom (2020s-Present). In recent years, large language models like GPT-3 and GPT-4 have demonstrated the potential of AI in understanding and generating human-like text. These models are trained on vast amounts of data and can perform a wide range of tasks, from answering questions to writing essays. The AI boom has also seen the integration of AI into various industries, enhancing productivity, and creating new opportunities.

At its core, AI-based programming involves creating algorithms that enable machines to perform tasks that typically require human intelligence. This includes learning from data, recognizing patterns, making decisions, and improving over time. The essence of AI lies in its ability to adapt and evolve, making it a powerful tool for solving complex problems.

The impact of AI-based programming is profound and far-reaching. In healthcare, AI is used for diagnosing diseases and personalizing treatment plans. In finance, it helps in fraud detection and algorithmic trading. Autonomous vehicles, powered by AI, promise to revolutionize transportation. Moreover, AI-driven innovations in natural language processing are transforming how we interact with technology, making it more intuitive and accessible. But artificial intelligence is also bringing breakthroughs in the programming of geoinformation systems. Although software is becoming more and more specialised, there will always be tasks for which there are no built-in functions [1] [2].

Geospatial Artificial Intelligence is an exciting and growing field that combines GIS (Geographic Information Systems) with Artificial Intelligence (AI) and Machine Learning (ML) for spatial data analysis, pattern recognition and various applications such as urban planning, environmental monitoring, disaster management and others. In the 2010's in the field of GIS and Remote Sensing an emergence of deep learning technologies, which are capable of much more accurate and complex analyses than traditional machine learning can be seen [3, 4].

Between 2015-2020: AI is increasingly being used for predictive analysis of geographic data, for example in weather, earthquakes, and other disaster forecasting. The use of CNNs for satellite image analysis and automated land use recognition is spreading. GeoAI emerged prominently when the integration of spatial data and artificial intelligence technologies facilitated the development of tools designed to address complex issues, including urban planning, crime prediction, climate change monitoring, and disaster management [5,6,7].

The evolution of AI-based programming reflects a journey from symbolic reasoning to data-driven learning, culminating in the development of sophisticated models that can perform a wide array of tasks. As AI continues to advance, its potential to drive innovation and solve real-world problems grows, heralding a future where intelligent systems are integral to our daily lives.

II. RESEARCH METHODOLOGY

The aim of the study is to code visual lisp programs for use in AutoCAD software with Microsoft Copilot. The proposed codes are imported each time and then run in AutoCAD software. The iteration process continues until the code is working properly. The test measures the development time and the number of iterations. During the study, interactions are recorded and typical errors are evaluated. The cause of the failure is investigated with both artificial and human intelligence.

Three different but interdependent pieces of code are generated. In geoinformatics and land surveying tasks, it is often possible to make edits based on point numbers and point codes. The first program has to generate the vector file based on the given points. In engineering practice, it is important that labels and dimensions are placed on the map. The second program should perform the area calculation and place the area data as labels in the centroid of the polygon. The third script should perform the scaling automatically according to the specified conditions. The aim is that all the drawing elements and captions that appear on the screen are created by the commands.

Task 1: Create a visual lisp program code for use in AutoCAD Civil 3D software as follows:

- If not, create a room layer that is bright green in color.
- Request a file containing the points of a polyline: point number, x coordinate, y coordinate in order, separated by commas.
- Connect the points in the room layer to create a closed polygon. So, connect the points in order by the score and finally connect the last point with the highest score to the first one.
- Do not use built-in functions!

Task 2: Create a visual lisp program code for use in AutoCAD Civil 3D software as follows:

- Calculate the area of polygons in the "room" layer based on the coordinates of the breakpoints, without built-in functions.
- Calculate the centroid of the polygons without built-in function and print it out.
- Create a yellow layer called "area".
- Insert the area data into the area layer as a label at the geometric centroid of the polygons.
- The height of the inscription should be 0,2. Add "m^2" at the end of the caption.

Task 3: Create a visual lisp program code for use in AutoCAD Civil 3D software as follows:

- Create a light blue layer.
- Scale the closed polygons in the room layer in the x and y directions:
- Each line segment must have its size written in the middle of the line in font size 0.2.
- In accordance with engineering practice, the inscriptions should face in two directions: horizontal inscriptions in the x-axis and vertical inscriptions in the y-axis. For display, use an offset so that the lettering does not obscure the line.

III. RESULTS

The code was written using Microsoft Copilot and Chat GPT Open AI models. The first task caught both AI models. The number of iterations reached the maximum, but the proposed code did not achieve the desired goal in either case. The majority of error messages (e.g. bad argument type, bad argument value, no function definition) are related to the reading of coordinates. For reading, prefix built-in functions (vl-string-split, ssget, read-line, vl-file-syst, split-string) were recommended, which did not work, even though the structure of the read file was modified.

At the end of the iterations, polygons were created that were linear, had more points than the definition, and the coordinates did not match (Fig. 1). The task was finally executed by a built-in script (Fig. 2).

The second task was successfully implemented by the Microcoft Copilot model. The result took 24 iterations, but finally the desired line of code was achieved. Most of the error messages are related to the centroid calculation. The centroid coordinates were calculated outside the polygons in negative range. Once the centroid was correctly determined, the area was easily displayed and formatted (Fig. 3).

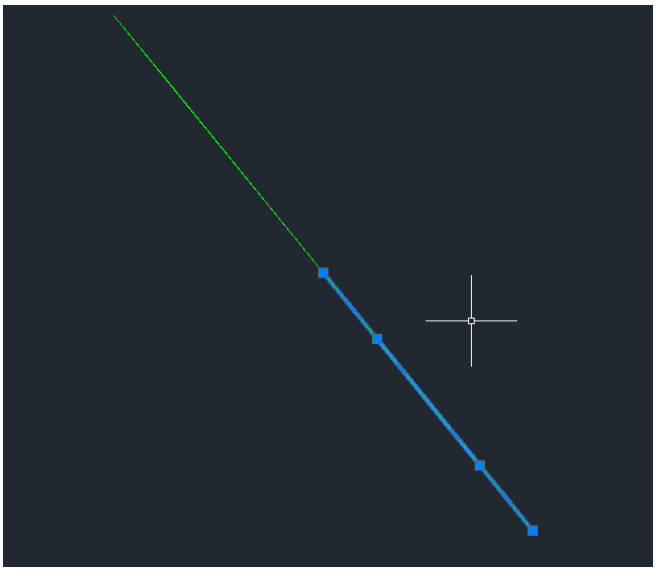


Fig. 1. Unsuccessful result of polygon construction

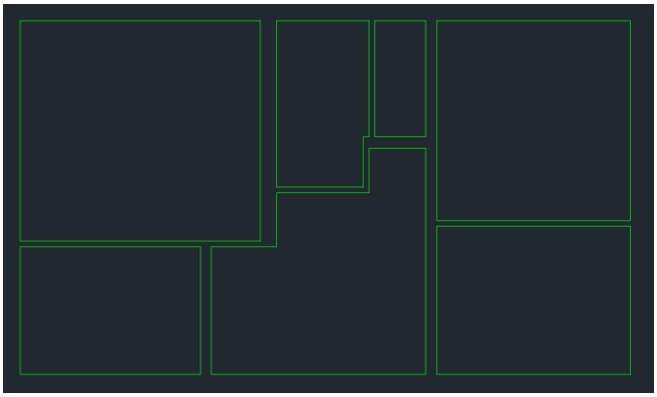


Fig. 2. Successful polygon construction with built-in script (PLINE)

The steps of the resulting solution without the specific lisp code:

1. 1. Setting up the Document and ModelSpace:
    - Get the active AutoCAD document.
    - Access the ModelSpace of the document.

    2. Creating the “area” layer if it doesn’t exist:
    - Check if the “area” layer exists.
    - If not, create the “area” layer and set its color to yellow.

    3. Defining the function to calculate area and centroid

    4. Calculating the angle of a point relative to the centroid.

    5. Calculating the centroid of the points:
    - Initialize variables for the sum of x and y coordinates.
    - Calculate the centroid by averaging the coordinates.

    6. Sorting points in clockwise order: Sort the points based on their angle relative to the centroid.

    7. Calculating the area and centroid:
    - Initialize variables for area and centroid coordinates.
    - Calculate the area using the shoelace formula.
    - Calculate the centroid coordinates.

    8. Iterating through all polylines in the “room” layer:
    - Loop through all entities in the ModelSpace.
    - Check if the entity is a polyline in the “room” layer.

    9. Handling points and coordinates:
    - Initialize a list for points.
    - Extract coordinates from the polyline.
    - Convert coordinates into points.

    10. Calculating and displaying the area and centroid:
    - Call the function to calculate the area and centroid.
    - If the result is valid, display the area and centroid.
    - Create a text entity to display the area in the drawing.

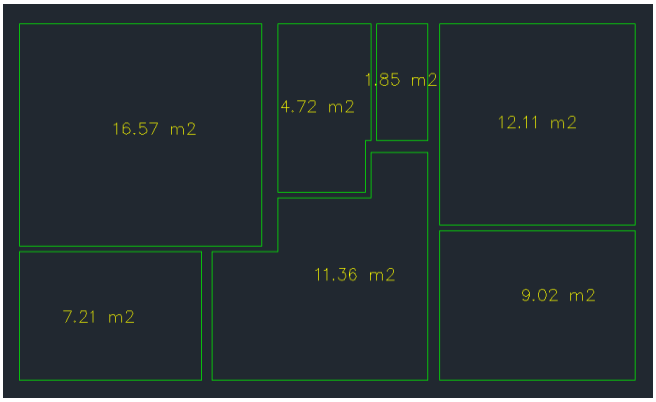


Fig. 3. Inscription of polygons by running the application developed by AI

The third task was solved by Chat GPT in 7 steps. During development, similar problems to the second task were encountered. Several built-in functions (e.g. VLAX-CURVE-GETPOINTS, Unknown command) were not available in AutoCAD Civil 3D 2023. The steps of the resulting solution without the specific lisp code:

1. Setting layers and colors:
- Define the name of the layer to be examined.
- Define the name of the dimension layer.
- Set the color to light blue (AutoCAD color code: 5).

2. Creating the dimension layer if it doesn’t exist:
- Check if the dimension layer exists.
- If not, create the new layer and set its color to light blue.

3. Searching for polygons on the specified layer:
- Select closed polygons on the specified layer.

4. Processing and dimensioning the polygons:
- Count the entities.
- Read each entity and convert it to a VLISP object.
- Get the number of vertices of the polygon.
- Dimension each line segment of the polygon.
- Calculate the midpoint and distance of each segment.
- Calculate the angle of the line segment.
- Calculate the perpendicular direction (90-degree offset inside the polygon).
- Apply the offset to the midpoint.

The first usable result is shown in Figure 4.

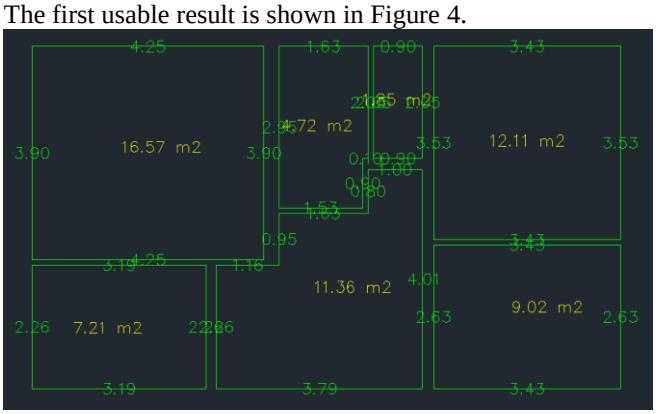


Fig. 4. Scaling of polygon elements with AI support – version 1

The AI has already made it easy to rotate the subtitle in the right direction (Fig. 5.):
1. Calculate the angle: using the two points (p1 and p2), calculate the angle of the line using the atan function, which gives the angle in radians based on the difference between the x and y coordinates of the line.
2. Text rotation: the text is rotated using the vla-put-Rotation method, where the angle of rotation is parallel to the line.

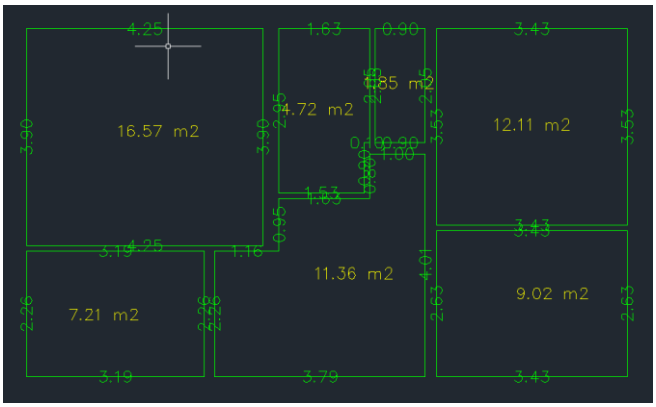


Fig. 5. Scaling of polygon elements with AI support – version 2

Figure 6. shows the solution produced by the code adopted as final.

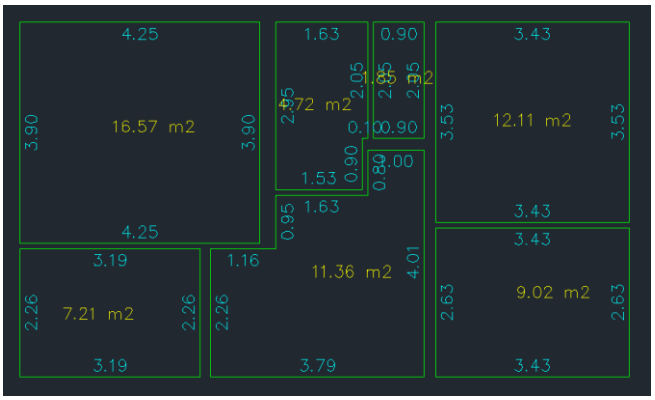


Fig. 6. Scaling of polygon elements with AI support – version 3

The development experience is presented in the Summary chapter.

IV. Summary

As literature shows that problem-solving methods proved that the application of artificial intelligence enables greater automation across various domains of engineering, allowing engineers to dedicate more time to solving creative problems. Moreover, the design processes can be expedited and simplified with the assistance of AI, thereby reducing the number of mechanical tasks involved.

As demonstrated by the presented examples, not all tasks can yet be executed seamlessly with the assistance of artificial intelligence. Coordinate reading proved to be problematic in both tested tools; however, these issues will likely be resolved in the future.

In this study, we tested the programming capabilities of AI for AutoCAD. It is conceivable that artificial intelligence may not possess as deep an understanding in this domain compared to other software applications. Additionally, it is possible that AI has not been as extensively trained to address engineering problems as it has been for solving more routine tasks.

As the presented task approved that during the execution of the presented problems, we observed that achieving an optimal solution required multiple iterations, sometimes up to 24. Furthermore, reaching a perfect solution necessitated a foundation of basic programming knowledge.

In summary, we observed the following advantages and disadvantages during the execution of the tasks.

The goal of solving the task without manual editing and subtitling was achieved. AI has many advantages:

- it does not make syntactic errors,
- it gives fast results,
- it documents the development flow and code well,

However, the presented application development also pointed out the weaknesses of AI, which are:

- it wants to use built-in functions even when asked to do the opposite,
- it repeatedly submits the same suggestion (e.g. when you ask it what the rewrite in the code is, it admits that it is nothing),
- puts checks in the code only on request or after 20-30 tries.

The conclusions drawn from the extension are:

- You have to be very good at rewriting the task. All information useful for coding should be provided.
- Professional terminology should be used, as it allows you to draw from relevant context.
- A check should be requested, as the validity of the result is not checked.
- Preferably, you should communicate in a formulaic world language, because language engines are not yet flawless.
- Provide a solution in all cases, even if you do not have the necessary information

Overall, AI-enabled programming requires basic programming skills. It took many iterations to get to the desired result, which was slower than expected.

References

[1] Kyung-Wan, S. - Young-Ha, S. - Jae-Hong, Y. - Dong-Cheon, L., Generating CityJSON-based Point Cloud Training Dataset for 3D Deep Learning, Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and CartographyVolume 40, Issue 6, Pages 487 – 500 (2022)

[2] László, G. - Molnár, G. P. -Bor, G. -Tóth, Z., Lever arm calculating algorithm for efficient and precise vector determinations, Czech Journal of Civil Engineering 9 : 1 pp. 35-41. , 7 p. (2023)

[3] Janowicz, K., Gao, S., McKenzie, G., Hu, Y., & Bhaduri, B. (2019). GeoAI: spatially explicit artificial intelligence techniques for geographic knowledge discovery and beyond. International Journal of Geographical Information Science, 34(4), 625–636. https://doi.org/10.1080/13658816.2019.1684500

[4] Györök, Gy. – Beszédes, B., Fault-tolerant Software Solutions in Microcontroller Based Systems. In: Orosz, Gábor Tamás (szerk.) AIS 2017 - 12th International Symposium on Applied Informatics and Related Areas organized in the frame of Hungarian Science Festival 2017 by Óbuda University : Proceedings. Székesfehérvár, Magyarország : Óbudai Egyetem (2017) 204 p. pp. 7-12. , 6 p.

[5] Györök, Gy. – Beszédes, B., Duplicated Control Unit Based Embedded Fault-masking Systems. In: IEEE - Szakál, Anikó; IEEE (szerk.) 2017 IEEE 15th International Symposium on Intelligent Systems and Informatics (SISY) : 15th IEEE International Symposium on Intelligent Systems and Informatics. New York, Amerikai Egyesült Államok : IEEE (2017) pp. 283-288. Paper: 8080567 , 6 p.

[6] Györök, Gy. – Beszédes, B., Fault tolerant power supply systems. In: Orosz, Gábor Tamás (szerk.) 11th International Symposium on Applied Informatics and Related Areas (AIS 2016). Székesfehérvár, Magyarország : Óbudai Egyetem (2016) pp. 68-73. , 6 p.

[7] Pengyuan Liu, Filip Biljecki,A review of spatially-explicit GeoAI applications in Urban Geography, International Journal of Applied Earth Observation and Geoinformation, Volume 112, 2022, 102936, ISSN 1569-8432,https://doi.org/10.1016/j.jag.2022.102936.

Remote Sensing Based Soil Moisture Assessment: Contributions to WREN project

Malgorzata Verőné Wojtaszek
Alba Regia Faculty
Óbuda University
Székesfehérvár, Hungary
wojtaszek.malgorzata@amk.uni-obuda.hu

Bence Ádám
Alba Regia Faculty
Óbuda University
Székesfehérvár, Hungary
adambence1989@gmail.com

Valéria Balázsik
Alba Regia Faculty
Óbuda University
Székesfehérvár, Hungary
balazsik.valeria@amk.uni-obuda.hu

Balázs Böröcz
Alba Regia Faculty
Óbuda University
Székesfehérvár, Hungary
borocz.balazs@amk.uni-obuda.hu

Máté Gfellner
Alba Regia Faculty
Óbuda University
Székesfehérvár, Hungary
gfellner.mate@stud.uni-obuda.hu

János Katona
Alba Regia Faculty
Óbuda University
Székesfehérvár, Hungary
dr.katona.janos@amk.uni-obuda.hu

Gábor Molnár
Alba Regia Faculty
Óbuda University
Székesfehérvár, Hungary
molnar.gabor@amk.uni-obuda.hu

Andrea Pődör
Alba Regia Faculty
Óbuda University
Székesfehérvár, Hungary
podor.andrea@amk.uni-obuda.hu

Gábor Nagy
Alba Regia Faculty
Óbuda University
Székesfehérvár, Hungary
nagy.gabor@amk.uni-obuda.hu

***Abstract*—The aim of the WREN (Water Resources in Efficient Networks) Project to develop a country-wide drought monitoring and forecasting system in Hungary, that combines the on-site soil moisture data of sparsely situated drought monitoring stations of the General Directorate of Water Management (OVF), and medium resolution satellite imagery.**

For modelling the spatio-temporal variations of the soil moisture a spatial database was generated. This database consists of high spatial resolution topographic, pedological and on-site soil moisture measurement data for sample areas, and low spatial resolution meteorological and soil moisture data for the last two years for the area of Hungary. Additionally, the SENTINEL-2 satellite imagery database provides a middle spatial (10m pixel size) and middle temporal (2-3 scenes weekly) resolution input data on national scale. For the sample area, a high resolutions Digital Surface Model (DSM) was generated, and used to analyze the relationship between the derivatives of the DSM and the soil moisture. For analyzing the relationship between the remotely sensed data and soil moisture, numerous indices (calculated from combining different frequency bands of satellite imagery) were generated and analyzed. Of the analyzed models, the two variable OPTRAM model provided best results with R¿90 correlations.

These results are crucial for the utilization of the recently built and launched WREN-1 satellite, whose imagery is supposed to fill the data gaps of the SENTINEL-2 satellite data.

***Index Terms*—Soil Moisture, Digital Elevation Model (DEM), indices, OPTRAM, SENTINEL-2**

I. Introduction

In addition to meteorological data (amount, timing and intensity of precipitation), other factors, such as soil and topography, also influence the occurrence of drought. The influence of topography on meteorology and the hydrological and temperature regime of soils cannot be neglected. Topography directly or indirectly controls the spatial distribution of physical, chemical and biological soil properties.

Previous research has already demonstrated the relationship between slope steepness and exposure and soil moisture. The infiltration dynamics of selected rainfall events at different latitudes and vegetation belts have been investigated, showing a clear correlation between both rainfall infiltration and evapotranspiration with these topographic features [1].

Some of the rainfall infiltrates into the soil. The resulting soil moisture is stored in the soil, partly by adhering to the soil particles and partly by the gaps between the particles. Plants can use this water, which is contained in the gaps between the soil particles. The water in the soil may be bound, which plants cannot or can only partially use, or free. Between waters include chemically between waters, biologically between waters, and waters bound by adsorption forces. Free waters include gravity water, groundwater and capillary water. [2] Because the distribution of precipitation is uneven, the soil must have the ability to bind water against gravity, but not to the extent that it cannot be absorbed by plants. The soil stores water in its pores. The porosity of soil can be described by its bulk density (how much mass per unit volume of soil). The total pores affect the water conductivity and aeration of the soil. It also depends on grain composition. The plant needs water and air, not pores. Soil with good water management can provide both water and air. Water is bound to the colloidal grains of the soil because of its dipole nature. In the immediate vicinity of the grains, the colloid binds water so strongly that

it cannot be taken up by the plant. As you move away from the colloid surface, the binding strength decreases. Eventually you reach a binding strength that loosely binds the water, meaning that gravity causes the water to be pulled away from the colloid. The binding numbers of different soils:

- sand 25-30,
- sandy loam 31-37,
- loam 38-42,
- clay loam 43-50,
- clay 51-60,
- heavy clay 61-80.

Water rises between the grains due to capillary action. This property of the soil is called capillary rise. Capillary rise is the result of the combined action of surface tension of water and adhesion of soil particles and water molecules. The rate and height of capillary rise depends on the size of the gaps and the swelling of the soil colloids. Overall, the better the soil structure, the more water it can hold. The deeper the topsoil, the more water it can store.[3] [4]

In Hungary the OVF system is responsible for drought monitoring. The system can predict water scarcity based on soil moisture data regularly collected from depths of 10, 20, 30, 45, 60 and 70 cm. Although the number of stations is high, 112, and they cover the entire country, precise field-level information is difficult to calculate from this data, and interpolation yields relatively inaccurate results. However, the data from the stations closest to the sample area can still be useful during validation.

The objective of the WREN project is to improve the current field drought monitoring system with higher temporal and spatial resolution data, including digital elevation and surface models [5].

II. Data, study area

A. Study area

The study area is situated in the northern region of Hungary, near Csolnok. Due to its topographic features, parts of the study area exhibit poor productivity as a result of surface erosion in several locations. The area is characterized by brown forest soil, with an organic matter content ranging from 1.00% to 2.00%. Soil sample results from 5-hectare zones were available for analysis. The most important soil properties for water management include physical characteristics (clay, loam, sand content) and organic matter content. One key physical property of soils is the ”Arany-féle” index, which is used to infer soil physical properties. This index (symbol KA) indicates how many cubic centimeters of water per 100 grams of air-dry soil can be absorbed up to the soil’s upper limit of plasticity. The Csolnok field (73 block) is characterized by sandy loam (KA: 31-37) and loam (KA: 38-42). The values measured in the laboratory (KA) range from 33 to 39. The organic matter content (H%) varies between 1.37 and 1.9%. However, the representativeness of the sample areas is not satisfactory for all zones. The soil samples taken from multiple points reflects only the heterogeneity of the recorded data, not the true characteristics of the area. See also Fig. 4 and the Chapter II-C.

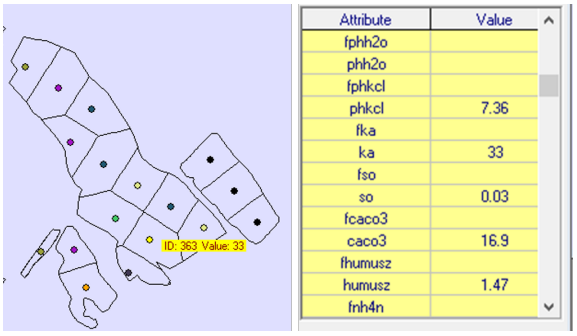


Fig. 1. The boundaries of the soil sampling zones (5 ha) and soil properties based on laboratory analysis.

The texture, depth, colour, chemical and physical properties, structure, and the arrangement of its horizons all characterize a soil and determine its agricultural, watermanagement and other uses. Soil vertical profiles and soil samples taken at different depths are an important source of data for sustainable use. A section of soil from the surface through its various layers down to the parent material is known as the soil profile. These layers are grouped into three main categories: A (eluvial region, zone of maximum leaching), B (illuvial region, zone of deposition of clay), and C (mineral mass). Soil erosion and other forms of degradation negatively impact the soil, primarily leading to a reduction in the thickness of the topsoil layer and a deterioration of its structure. The following figure (Fig. 2) shows a profile of areas with different slopes and erosion. The soil profiles are from areas used as pasture: 1. Brown forest soil: The degree of connectivity (KA) varies by depth (KA: 52, 41, 36). 2. Alluvial and sloping hardpan (soil material is present within 100 cm of the surface due to the accumulation of eroded soil material): The KA value varies with depth (KA: 54, 38, 32). 3. Brown forest soil: The degree of cohesion (KA) also varies by depth (KA: 64, 40, 28).

B. In situ measurements

During the first two years of the project, soil moisture was measured four times per year in the study areas. The timing of these measurements was determined by the crop growth stages and weather conditions. Field measurements were conducted during dry periods (drought) and after days with heavy rainfall. Soil moisture was measured in the root zone at depths of 10 and 20 cm. The spatial distribution of measurement points was aligned based on the heterogeneity of the agricultural field. The variability of the area was mapped by analyzing data from several sources. Satellite data classification results were supplemented with laboratory analyses of soil samples and field topographic features. However, it is important to note that the average soil sampling per 5 hectares may not always reflect the properties of specific sampling zones (Fig. 3). For example, the zone in the south-western part of the field includes both low- and high-fertility areas, meaning that

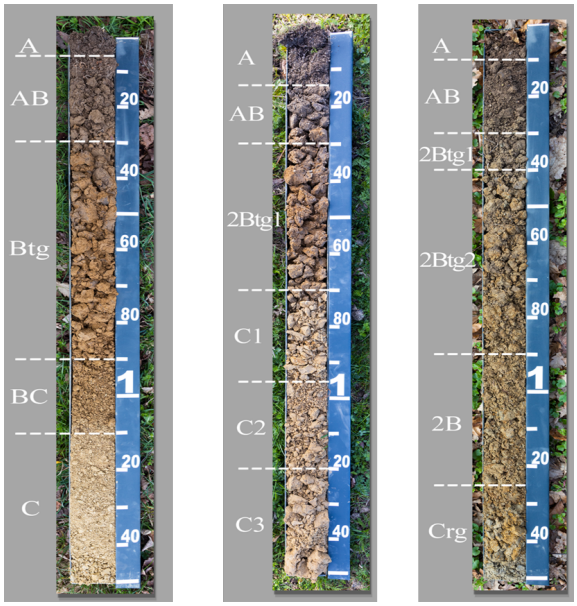


Fig. 2. Examples of soil profiles

a single soil sample cannot be considered representative in such cases. Soil moisture measurements were therefore taken in both parts of this area.

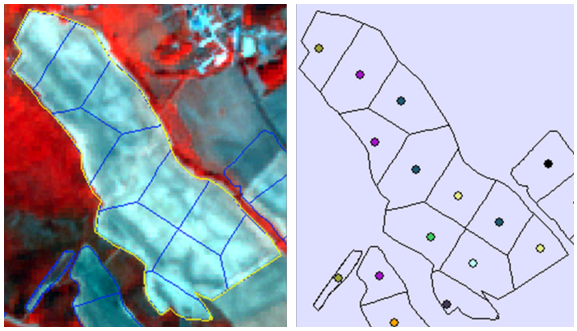


Fig. 3. The boundaries of the soil sampling zones (5 ha) overlaid on the satellite image (Sentinel2) and a map with the locations of the average sampling points within each zone.

Two sensors were used to determine soil moisture: the WaterScout SM100 Soil Moisture Sensor with a FieldScout Soil Sensor Reader, and the Spectrum FieldScout TDR 350 Soil Moisture Meter with Bluetooth & GPS connectivity. Both devices were calibrated for mineral soils, and the Volumetric Water Content (VWC%) of the soil was measured. To navigate the field and return to the same measurement points, we used a mobile GPS application with an accuracy of 1-2 meters. In this paper, the measurement results from a single date (1st of June, 2023) are presented as an example.

Prior to fieldwork, Sentinel2 multispectral imagery (VIS, NIR, SWIR bands) was classified into three different categories of productivity using a segment-based classification method. In the next step, each thematic class was divided into cells based on homogeneity. The calculation of the homogeneity criterion was based on mean values and standard deviation for each band. The size of each test cell was 4x4 pixels, with each pixel measuring 10 meters. In the field, after identifying the cell, several replicate soil moisture measurements were taken. A total of 367 measurements were collected, and the average measurement per cell was calculated before statistical analysis. Soil moisture values ranged from 6 VWC% to 45 VWC% (Fig. 4). High soil moisture values were typical in the high-productivity zones, which are located in the lower areas of the landscape. In contrast, low moisture values were measured in the hilly, eroded regions.

Fig. 4. Points of measurements were taken using the Spectrum FieldScout TDR 350 sensor, along with the average and spatial distribution of measurements for a test cell. Each cell consists of 4x4 pixels, with each pixel measuring 10 meters.

Date	Sensor type	Number of measured points	Moisture content - min	Moisture content – max	14-day precipitation amount (mm)	Drought index (HDI)	OVF SM 10cm	OVF SM 20cm
01-06-2023	SM 100	6	4,5	32,6	9,6	1,23332	26,00	26,08
01-06-2023	TDR 350	367	6,2	45,29				

TABLE I
SOIL MOISTURE MEASUREMENTS RESULTS AND OVF CSOLNOK STATION DATA

The table summarizes the soil moisture measurements, the 14-day precipitation amounts prior to the measurement day, and the soil moisture data recorded at two different depths at the OVF measuring station in Csolnok. A total of 367 point measurements were recorded during the measurement period.

C. Remote sensing data

In this study, we utilized multispectral imagery from the European Space Agency's Sentinel2 satellite, acquired via the ESA Sentinel Scientific Data Hub. Sentinel2 provides high spatial resolution, ranging from 10 to 60 meters, across 13 spectral bands that cover the visible (VIS), near-infrared (NIR), and shortwave infrared (SWIR) regions of the electromagnetic spectrum, with a temporal resolution of approximately 5 days. To ensure data quality, only cloud-free images were initially considered. After a preliminary interpretation, additional images were selected for further processing, taking into account meteorological factors such as the amount and intensity of precipitation, as well as the study area's coverage and fieldwork data. Image processing was performed using the VIS, NIR, and SWIR bands, with spatial resolutions of

10 and 20 meters. Remote sensing data have been analyzed from several aspects. One application involves mapping the spatial and temporal variability within the sample area to aid in planning soil sampling points. Another involves deriving indices, indicators, and models from selected spectral bands (VIS, NIR, SWIR), which form the foundation of a remote sensing-based drought monitoring system [6]. Before selecting the indicators, regression analysis was performed to assess the strength of the relationships between the reference data (field measurements) and the indices and models derived from satellite data. The results from field measurements, indices derived from Sentinel2 satellite data, and statistical analyses are being used to further develop the soil moisture model (OPTRAM).

D. Digital elevation models

Elevation date were obtained globally available remotely sensed arcsecond resolution SRTM raster Digital Elevation Model and a 10m resolution vector topography model of the sample area derived from the elevation contour lines of topographic maps. High-resolution aerial photography and laser scanning of the selected Csolnok 73 field in the sample area was also carried out using UAV technology, resulting a 1m resolution DEM. The resolution of these data are comparable with the planned 17m resolution of the WREN satellite optical and short wave infrared images.

The SRTM model was resampled to EOV (Standard Hungarian Projection, EPSG=23700) with a 25 meters GRID. The 10 meter resolution file was provided in this projection.

For the study, a 20 m field resolution digital model of the 65 ha Csolnok 73 plot was chosen, taking into account the remote sensing data, and the results of soil moisture measurements taken on 1st of June, 2023 at 13 on-site sampling points (Fig. 5).

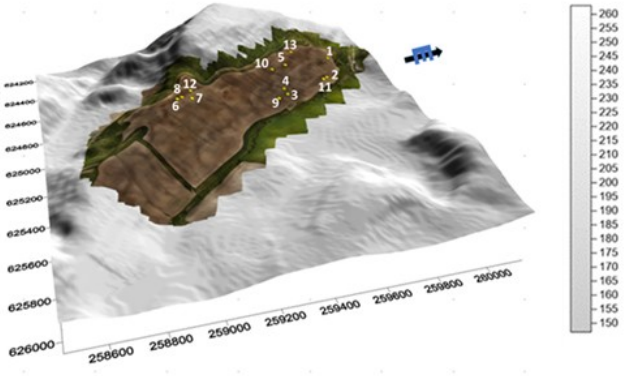


Fig. 5. Ortophoto of Csolnok 73 field on Digital Elevation Model with soil moisture sampling points of 1st of June, 2023.

III. Methods and Results

A. Analysis of elevation models

From the elevation values of the digital elevation and surface models, additional terrain characteristics, slope categories and aspect values can be derived, which have a strong influence on the water absorption and water storage capacity of soils [7]. The resulting quantities can then be further analysed to determine the relationship between terrain characteristics and soil moisture and to take them into account in drought risk assessment.

In the sample field, the slope categories are shown according to the category boundaries used in agriculture (Fig. 6).

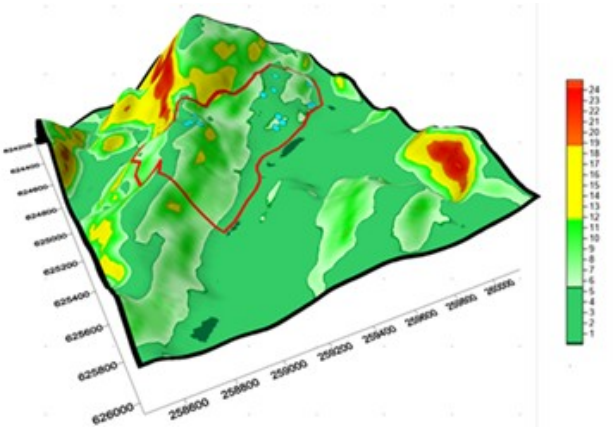


Fig. 6. Slope categories of Csolnok 73 field, according the slope nomenclature used in the agriculture

measured point	slope category	aspect	soil moisture (V/V%)
1	II	N	14.7
2	I	N	33.9
3	II	N	15.9
4	II	N	17.1
5	I	N	31.5
6	I	N	32.1
7	II	N	25.5
8	I	N	33.2
9	I	N	32.2
10	I	N	31.9
11	I	N	26.7
12	I	N	41.0
13	I	N	35.5

TABLE II
THE SLOPE CATEGORY [7], ASPECT AND MEASURED SOIL MOISTURE AT 10 CM DEPTH FOR THE SAMPLING POINTS ON 1ST OF JUNE, 2023.

The Table II includes the number of soil sampling points, the slope category and aspect of the point locations, and the soil moisture measured in the field on 1st of June 2023 at a depth of 10 cm.

B. Relationships between soil moisture and topographical indicators

Only two small spots on the Figure 6 show the steeper areas falling into slope category III, with the largest proportion of the slope values on the field falling into categories I and II according to the FAO's classification [7]. The on site measurement sites were selected on the basis of Sentinel space imagery, taking into account other criteria, so that the location of the 13 measurement points is random with respect to the topography. Based on the 20 m resolution model, the field is

mostly oriented N-S and all measurement points are located in this part of the field, so the relationship with soil moisture values by aspect cannot be investigated through this example. For the spatial representations, the height is exaggerated by a factor of 1.5.

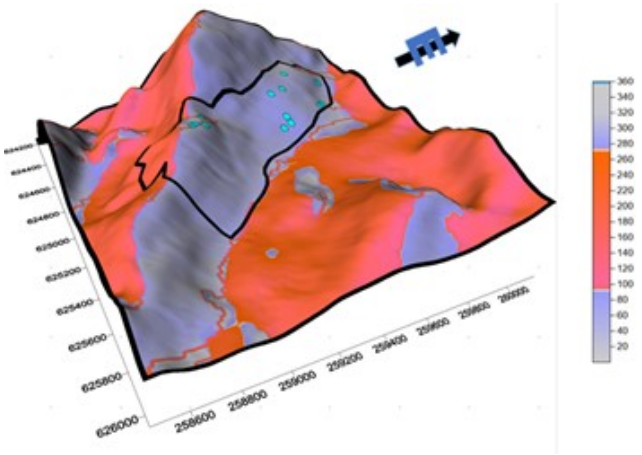


Fig. 7. Slope aspect of Csolnok 73 field. As the field is situated on the NE slope, its aspect is almost uniform

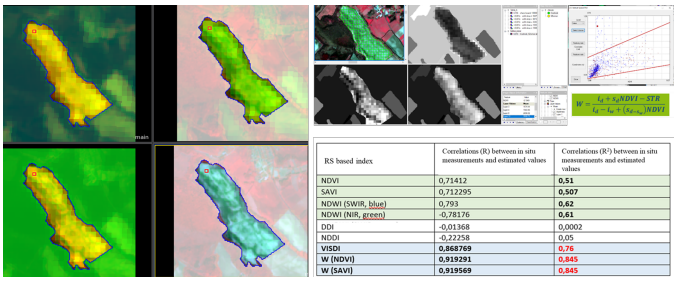
	Index	Megnevezés	Index képlete
1.	NDVI	Normalized Difference Vegetation Index	NDVI = (NIR − red) / (NIR + red)
2.	NDRE	Normalized Difference Red Edge Index	NDRE = (NIR – red edge) / (NIR + red edge) *B6*
3.	SAVI	Soil-Adjusted Vegetation Index	SAVI = (1+L)*((NIR-red)/(NIR+red))
4.	OSAVI	Optimized Soil Adjusted Vegetation Inde	OSAVI = NIR-red/NIR+red+K K=0,16 *(korrekciós tényezőjének)*
5.	NDWI (SWIR, blue)	Normalized Difference Water Index (Modify)	NDWI = (Blue - NIR)/(Blue + NIR)
6.	NDWI (NIR, green)	Normalized Difference Water Index	NDWI = (Green - NIR)/(Green + NIR)
7.	MNDWI (SWIR, green)	Modify Normalized Difference Water Index	MNDWI = (Green - SWIR)/(Green + SWIR)
8.	NDMI, Nedvesség Index	Normalized Difference Moisture Index	NDM I= NIR−SWIR/NIR+SWIR
9.	DDI *DVI, DWI*	Distance Drought Index *DVI (Difference Vegetation Index)* *DWI (Difference Water Index)*	DDI = DVI – DWI *DVI = NIR − red* *DWI = NIR − SWIR*
10	NDDI	Normalized Difference Drought Index	NDDI = (NDVI − NDWI) / (NDVI + NDWI) *NDWI= (Green - NIR)/(Green + NIR)*
11	VISDI	Visible and Shortwave Infrared Drought Index	VSDI = 1 − [(SWIR − BLUE) + (RED − BLUE)]
12	W *(NDVI)*	OPTRAM	W=i_d+s_d NDVI_- STR / i_d – i_w+(s_d-s_w)NDVI
13	W *(SAVI)*	OPTRAM	W=i_d+s_d NDVI_- STR / i_d – i_w+(s_d-s_w)NDVI

TABLE III
THE INDICES CONSIDERED IN THE STUDY AND THE METHOD OF THEIR CALCULATION

C. Remote sensing based soil moisture modelling

Vegetation and moisture indices (index images) are relatively easy to generate from multispectral images and can independently provide valuable information on the condition of plants or soil. NDVI and other vegetation indices have been effectively used for decades to map vegetation and track changes over time.

Research has demonstrated that water or moisture indices are well-suited for drought mapping, as there is a strong correlation between chlorophyll levels and vegetation moisture content. Further analysis of the relationship between water indices (e.g., NDWI) and vegetation indices (e.g., NDVI) can help investigate how vegetation responds to drought and quantify biomass changes due to water deficits (e.g. [8]).



RS based index	Correlations (R) between in situ measurements and estimated values	Correlations (R^2) between in situ measurements and estimated values
NDVI	0,71412	**0,51**
SAVI	0,712295	**0,507**
NDWI (SWIR, blue)	0,793	**0,62**
NDWI (NIR, green)	-0,78176	**0,61**
DDI	-0,01368	0,0002
NDDI	-0,22258	0,05
VISDI	**0,868769**	**0,76**
W (NDVI)	**0,919291**	**0,845**
W (SAVI)	**0,919569**	**0,845**

Fig. 8. An example of index visualization includes NDWI, NDVI, W (NDVI), W (SAVI), and the OPTRAM model calculation process, as well as indicators showing the strongest correlations.

However, the strength of the correlation (R^2) between soil moisture and indices like NDVI or NDWI is relatively modest, typically ranging between 0.5 and 0.6. Index images can also serve as inputs for more complex indicators when studying specific phenomena [9]. A common example is the normalized differential drought index (NDDI), which is calculated using a combination of the normalized differential vegetation index (NDVI) and the normalized differential water index (NDWI)—specifically, as the difference between the two indices divided by the sum of their values. These indices can also be used as input data for models, such as the OPTRAM model.

The indices considered in this study are summarized in the Table III. They are calculated using visible (VIS), near-infrared (NIR), and shortwave infrared (SWIR) spectral bands. The selected indices from the literature were categorized based on the following criteria: simple indices, complex indices, and model-based indices. Statistical methods were employed to test the correlations between in situ measurements and indicators derived from remotely sensed data. These calculations were performed for all time periods corresponding to the field soil moisture measurements. In this paper, the measurements and analyses for June 2023 are presented (Fig. 8.). Remote sensing data analysis was confined to the study area, utilizing a hierarchical framework to delineate the field of interest at the super-object level. Within the study area, analysis was conducted at the sublevel, with each unit consisting of 4 pixels, allowing soil moisture values to be determined at 40-meter intervals. Following multi-level segmentation, we calculated the necessary characteristics for each study unit to enable index and model execution, including vegetation and water indices, as summarized in the table.

In the context of water stress models (OPTRAM), the normalized differential vegetation index (NDVI, SAVI) and shortwave infrared transformed reflectance (STR) were used as

input data. NDVI was computed using reflectance values from the red band (B4) and near-infrared band (B8), while STR was calculated using reflectance values from the shortwave infrared band (B11) (Fig. 8).

During the statistical analysis, index-based values and OPTRAM-derived surface soil moisture estimates were compared with ground truth soil moisture data. The indices and estimated moisture values (W) that showed significant correlations with in situ measurements (R^2 = 0.5 - 0.85) are presented in the figure. Based on the Pearson correlation coefficient, the strongest positive relationship was found between the OPTRAM model and ground truth variables (R = 0.9). These results confirm the applicability of the model.

D. Automated OPTRAM modelling

For the effective application of the OPTRAM model, an automated estimation of the parameters was investigated.

In this process, the data were processed using the Google Earth Engine (GEE) development environment, a cloud-based BigData geospatial environment [10].

This cloud-based environment allows the analysis of multiple satellite images at the same time, and the system's database contains the Sentinel-2 satellite images, eliminating the need to download and store satellite images.

The objective of this was to perform the calculation of NDVI and STR indices on Sentinel-2 satellite images, and sampling these values for the on-site soil moisture sampling sites. The data tables containing these data were joined with the soil moisture (W) data, forming a three column data table for a selected date.

Using the NDVI-STR-W data points the four coefficients of the nonlinear equation of the OPTRAM model were calculated using the least squares method [11].

The surface representing the two-variable function of the soil moisture of the OPTRAM model is shown on Fig. 9.

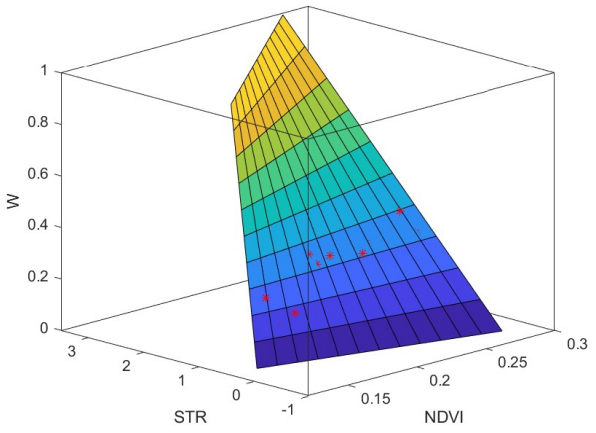


Fig. 9. The OPTRAM model of the soil moisture (W) as a function of the remotely sensed NDVI and STR indices. The surface represent the mathematical relationship. The coefficients of the fitted surface were calculated using the on-site measured soil moisture data (red stars)

Using the automated method, we found, that the standard deviation of calculated soil moisture (W) values are 3.25%, which is in good accordance with the results of the standard OPTRAM modelling.

IV. Conclusions

Analysing the effect of slope category on soil moisture, four points of the 13 measurements fall into the higher slope category II, and the soil moisture values measured at these sites, ranging from 14.7 to 25.5, are the four lowest among the measured results. The next highest soil moisture value in order of magnitude was measured at point 11 (26.7), which is just above the highest value in category II.

The other 8 measurement points with soil moisture values between 31.5 and 41.0 percent are all at on the moderate slopes of category I. Based on the presented results, a clear correlation between soil moisture and slope category can be concluded.

In the study, 13 different RS-based indicators were taken into consideration. The correlation between in situ measurements and estimated values was determined. According to the results of the statistical analysis (Pearson correlation coefficient), the strongest positive relationship was found between the OPTRAM model and ground truth variables (R = 0.9). These findings confirm the applicability of the model.

Acknowledgment

The authors gratefully acknowledge the support provided for this research by the GINOP_PLUSZ-2.1.1.-21-2022-00040 Project. The project is supported by Széchenyi Plan Plus, co-funded by the Government of Hungary and the European Union.

References

[1] S. Singh, "Understanding the role of slope aspect in shaping the vegetation attributes and soil properties in montane ecosystems," *Tropical Ecology*, vol. 59, no. 3, pp. 417–430, 2018.

[2] T. Keller, J. Arvidsson, and A. R. Dexter, "Soil structures produced by tillage as affected by soil water content and the physical quality of soil," *Soil and Tillage Research*, vol. 92, no. 1, pp. 45–52, 2007, ISSN: 0167-1987. DOI: https://doi.org/10.1016/j.still.2006.01.001. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S016719870600002X.

[3] M. Bittelli, "Measuring soil water content: A review," *HortTechnology hortte*, vol. 21, no. 3, pp. 293–300, 2011. DOI: 10.21273/HORTTECH.21.3.293. [Online]. Available: https://journals.ashs.org/horttech/view/journals/horttech/21/3/article-p293.xml.

[4] B. Tóth, A. Makó, and G. Tóth, "Role of soil properties in water retention characteristics of main hungarian soil types," *Journal of Central European Agriculture*, vol. 15, pp. 137–153, 2014. [Online]. Available: https://api.semanticscholar.org/CorpusID:83143968.

[5] M. Verőné Wojtaszek, G. Molnár, V. Balázsik, *et al.*, "Water resources in efficient networks (wren) project overview and presentation of the research results carried out during the first year," in *AIS 2023 - 18th International Symposium on Applied Informatics and Related Areas - Proceedings*, I. Petőné Csuka and G. Simon, Eds., Székesfehérvár, Hungary, 2023, pp. 90–98.

[6] M. D. Svoboda, B. A. Fuchs, *et al.*, *Handbook of drought indicators and indices*. World Meteorological Organization Geneva, Switzerland, 2016, vol. 2.

[7] Agriculture Organization of the United Nations. Soil Resources and Conservation Service and Agriculture Organization of the United Nations. Interdepartmental Working Group on Land Use Planning, *Guidelines for land-use planning*. Food & Agriculture Org., 1993, vol. 1.

[8] M. Sadeghi, S. B. Jones, and W. D. Philpot, "A linear physically-based model for remote sensing of soil moisture using short wave infrared bands," *Remote Sensing of Environment*, vol. 164, pp. 66–76, Jul. 2015. DOI: 10.1016/j.rse.2015.04.007.

[9] M. Verőné Wojtaszek, V. Szabó, J. Kauser, A. Kocsis, and L. Lippmann, "55. mapping of soil moisture variability within a field by the OPTRAM model," in *Precision agriculture '21*, Wageningen Academic Publishers, Jun. 2021. DOI: 10.3920/978-90-8686-916-9_55.

[10] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore, "Google earth engine: Planetary-scale geospatial analysis for everyone," *Remote Sensing of Environment*, vol. 202, pp. 18–27, 2017, Big Remotely Sensed Data: tools, applications and experiences, ISSN: 0034-4257. DOI: https://doi.org/10.1016/j.rse.2017.06.031. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0034425717302900.

[11] E. Mikhail and F. Ackermann, *Observations and Least Squares* (IEP series in civil engineering). IEP, 1976.

Comparison of different database management software in analyzing of drought monitoring system data.

Gábor Nagy*
*Óbuda University, Alba Regia Technical Faculty, Institute of Geoinformatics,
nagy.gabor@amk.uni-obuda.hu
ORCID ID: 0000-0001-9453-2291

***Abstract*—The OVF (Országis Vízügyi Főigazgatóság, the hungarian hydrological authorities) has installed a network of drought monitoring stations to measure of various properties (rainfall, temperature, soil water content in several depth, etc.) in more than hundred points. There is an API that provides the collected data via HTTP in a JSON file.**

This paper describes a relational database for the available drought monitoring data, and compare the implementation with different database management software.

***Index Terms*—database management software, drought monitoring, spatial databases**

I. Introduction

From the web server of the hungarian hydrological authorities (Országis Vízügyi Főigazgatóság, OVF) we can download a lot of drought related data. The authority operates a network of stations, which measures the basic meteorological properties (temperature and rainfall) and the water content and the temperature of the soil at various depth (10, 20, 30, 45, 60 and 75 centimeter) in hourly epochs.

The data is available by an published API. The appropriate HTTP requests can download the specifies data in a JSON dataset. I made several Python [1], [2] scripts for this task when we work in the WREN project [3]. The created scripts download the necessary data from the OVF server and write to a suitable file specified by the other researchers of the project.

I write one more program for downloading the whole dataset and storing in a database. This paper describes this database and compares two solution for the embedding database.

SQLite

DuckDB

II. The database scheme

OVF's drought dataset has three dimension: the station (the place), the type of the measure and the time of the measurement. The network has 121 stations which measure 15 measurements at hourly epochs.

The created database scheme has two tables. The "allomas" contains the id, name and position of a station, and the "adat" stores one station measurements in one epoch in the records.

The dataset from 2017 to 2024 September was written to an SQL file (a text file with SQL commands, that build the database). The size of this file is 588 MiB.

The head of the SQL file contains the DDL command, that create the two relations of the database.

```
-- Az OVF Aszálymonitoring redszer adatai
-- 2024.10.01 és 2024.10.15 között


CREATE TABLE allomas (
    allid integer PRIMARY KEY,
    allnev varchar(30) NOT NULL,
    eovY float,
    eovX float
);


CREATE TABLE adat (
    allid integer,
    idopt timestamp,
-- Levegöhmérséklet [°C]
    lhom     float,
-- Talajhömérséklet (10 cm) [°C]
    thom10   float,
-- Talajhömérséklet (20 cm) [°C]
    thom20   float,
-- Talajhömérséklet (30 cm) [°C]
    thom30   float,
-- Talajhömérséklet (45 cm) [°C]
    thom45   float,
-- Talajhömérséklet (60 cm) [°C]
    thom60   float,
-- Talajhömérséklet (75 cm) [°C]
    thom75   float,
-- Talajnedvesség (10 cm) [V/V %]
    tnedv10  float,
-- Talajnedvesség (20 cm) [V/V %]
    tnedv20  float,
-- Talajnedvesség (30 cm) [V/V %]
    tnedv30  float,
-- Talajnedvesség (45 cm) [V/V %]
    tnedv45  float,
-- Talajnedvesség (60 cm) [V/V %]
    tnedv60  float,
-- Talajnedvesség (75 cm) [V/V %]
```

```
    tnedv75  float,
-- Csapadek60 [mm]
    csap60p  float,
-- Relatív páratartalom [%]
   relpar   float,
PRIMARY KEY (allid, idopt) );
```

The followed part of the file contains the DML commands (INSERT INTO) in a transaction, which insert the data to the tables.

The database does not have geospatial extension, but there are coordinates in the table of the stations (allomas) and we can create the WKT representation of the points of the stations:

```
-- The drought monitoring stations
-- with WKT representation of the point
SELECT allid, allnev,
  'POINT('||eovY||' '||eovX||')' as wkt
FROM allomas;
```

If the database has geospatial extension (SpatiaLite or PostGIS [4], [5]), we can create the point geometry data:

```
-- The drought monitoring stations
-- with point geometry
SELECT allid, allnev,
  ST_Point(eovY, eovX, 23700) as geom
FROM allomas;
```

III. The database software's

I use two embedded database software for storing the collected data and make queries from this dataset. Both of those programs can run the SQL script created by the download data.

The SQLite [6], [7] is a well known open source embedded database library. We use this program in the education of our institute for learning the SQL.

The DuckDB [8] is an OLAP (online analytical processing) embedded database, unlike the SQLite which is an OLTP (Online transactional processing) solution. This software is recommended for work with large databases in the local computer.

Both of the databases use by DBeaver interface.

The size of the SQLite database file is 903 MiB, and the size of the DuckDB database file is 298 MiB. (The size of the SQLite file is 303% of the DuckDB file size)

IV. Method of comparison

I compare the performance of the introduces databases by three typical SQL queries. I run these queries ten times and calculated the average runtime in both of the databases.

The first query calculates the aggregated values in a year (2023) at each station of the network.

```
-- The aggregated values in 2023
-- at each station
SELECT allomas.allid,
       allomas.allnev,
       sum(csap60p),
       avg(relpar),
       avg(lhom),
       avg(thom10),
       avg(thom20),
       avg(thom30),
       avg(thom45),
       avg(thom60),
       avg(thom75),
       avg(tnedv10),
       avg(tnedv20),
       avg(tnedv30),
       avg(tnedv45),
       avg(tnedv60),
       avg(tnedv75)
FROM adat JOIN allomas
     ON adat.allid=allomas.allid
WHERE idopt
   between '2023-01-01 00:00:00.000'
   and '2023-12-31 23:59:00.000'
GROUP BY 1,2
ORDER BY 2;
```

The average runtime of the firs query is 0.882 second in the SQLite and 0.118 seconds in the DuckDB. The SQLite runtime is 747% of the DuckDB runtime.

The second query calculates the available values (the working stations) in each epoch.

```
-- The count of available measures
-- at each measurement in 2023
SELECT idopt,
     count(csap60p),
     count(relpar),
     count(lhom),
     count(thom10),
     count(thom20),
     count(thom30),
     count(thom45),
     count(thom60),
     count(thom75),
     count(tnedv10),
     count(tnedv20),
     count(tnedv30),
     count(tnedv45),
     count(tnedv60),
     count(tnedv75)
FROM adat
GROUP BY 1
ORDER BY 1;
```

The average runtime of the second query is 5 seconds in the SQLite and 0.478 seconds in the DuckDB. The SQLite runtime is 1046% of the DuckDB runtime.

The third query calculates the daily average of the temperature in every station where all of the 24 measurements is available.

```
--Daily average temperatures
--DuckDB version
--in SQLite use date(idopt)
--instead of strftime(idopt, '%Y-%m-%d')
SELECT allomas.allid,
       allomas.allnev,
```

```
       strftime(idopt, '%Y-%m-%d'),
       avg(lhom)
FROM adat JOIN allomas
  ON adat.allid=allomas.allid
GROUP BY 1,2,3
HAVING count(lhom)=24;
```

The two inspected database have different functions for timestamp. The SQLite query has a little different.

```
--Daily average temperatures
--SQLite version
--in DuckDB use strftime(idopt, '%Y-%m-%d')
--instead of date(idopt)
SELECT allomas.allid,
       allomas.allnev,
       date(idopt),
       avg(lhom)
FROM adat JOIN allomas
  ON adat.allid=allomas.allid
GROUP BY 1,2,3
HAVING count(lhom)=24;
```

The average runtime of the third query is 5.1 second in the SQLite and 0.508 seconds in the DuckDB. The SQLite runtime is 1004% of the DuckDB runtime.

The fourth query is similar than the third. This query calculates the daily sum of the rainfall in every station where all of the 24 measurements is available.

```
--Daily rainfall
--DuckDB version
--in SQLite use date(idopt)
--instead of strftime(idopt, '%Y-%m-%d')
SELECT allomas.allid,
       allomas.allnev,
       strftime(idopt, '%Y-%m-%d'),
       sum(csap60p)
FROM adat JOIN allomas
  ON adat.allid=allomas.allid
GROUP BY 1,2,3
HAVING count(csap60p)=24;
```

The average runtime of the third query is 5.8 second in the SQLite and 0.565 seconds in the DuckDB. The SQLite runtime is 1027% of the DuckDB runtime.

The fifth query calculates the count of the stations where all of the 24 measurements is available in each day.

```
--The count the stations,
--where the rainfall was measured all day
--DuckDB version
--in SQLite use date(idopt)
--instead of strftime(idopt, '%Y-%m-%d')
SELECT alldatcsap.datum, count(*) as meres
FROM (SELECT allid,
        strftime(idopt, '%Y-%m-%d') as datum
      FROM adat
      GROUP BY 1, 2
      HAVING count(csap60p)=24) as alldatcsap
GROUP BY datum
ORDER BY 1
```

The average runtime of the third query is 7.2 second in the SQLite and 0.457 seconds in the DuckDB. The SQLite runtime is 1574% of the DuckDB runtime.

The sixth query calculates the minimum, maximum and average air temperature in 2023 at each station.

```
-- The minimal, maximal
-- and average temperature
-- in 2023 at each station
SELECT allomas.allid,
       allomas.allnev,
       allomas.eovy,
       allomas.eovx,
       min(lhom) minho,
       max(lhom) maxho,
       avg(lhom) atlagho
FROM adat JOIN allomas
     ON adat.allid=allomas.allid
WHERE idopt
   between '2023-01-01 00:00:00.000'
   and '2023-12-31 23:59:00.000'
GROUP BY 1,2,3,4
ORDER BY 2;
```

The average runtime of the third query is 0.519 second in the SQLite and 0.069 seconds in the DuckDB. The SQLite runtime is 748% of the DuckDB runtime.

The seventh query contains the minimum and the maximum temperature in each stations with the name of the one ore more station where these value was measured.

```
-- Maximal and minimal air temperature
-- at each day with staion names
SELECT nr.idopt,
 (SELECT string_agg(allnev, ', ')
  FROM allomas
  WHERE nr.minhom
    in (SELECT lhom
        FROM adat
        WHERE allomas.allid=adat.allid
          AND nr.idopt=adat.idopt)) as minall,
  minhom,
 (SELECT string_agg(allnev, ', ')
  FROM allomas
  WHERE nr.maxhom
    in (SELECT lhom
        FROM adat
        WHERE allomas.allid=adat.allid
          AND nr.idopt=adat.idopt)) as maxall,
  maxhom
FROM (SELECT idopt,
       min(lhom) as minhom,
       max(lhom) as maxhom
      FROM adat
      GROUP BY 1) as nr
ORDER BY 1;
```

The average runtime of the third query is 64.3 second in the SQLite and 2 seconds in the DuckDB. The SQLite runtime is 3215% of the DuckDB runtime.

Table I. The runtimes of the test queries in seconds, and the ratio between the SQLite and DuckDB runtime

num	SQLite	DuckDB	ratio
1	0.882	0.118	7.47
2	5.000	0.478	10.46
3	5.100	0.508	10.04
4	5.800	0.565	10.27
5	7.2	0.457	15.74
6	0.519	0.069	7.48
7	64.3	2.000	32.15

V. Conclusion

The previous results are in the Table I.

The DuckDB queries are 7-15 times or even 30 times faster and the DuckDB database file size is a third than the SQLite.

The DuckDB seems more effective, but I also can do the queries by the SQLite only some seconds. The following work I probably use both of the databases.

In the future these test will run in databases servers (for example MySQL, PostgreSQL, etc.) and the test query set may be expanded.

References

[1] G. Van Rossum *et al.*, “Python Programming Language.” in *USENIX Annual Technical Conference*, vol. 41, 2007.

[2] T. E. Oliphant, “Python for scientific computing,” *Computing in Science & Engineering*, vol. 9, no. 3, 2007.

[3] V. W. Malgorzata, M. Gábor, B. Valéria, P. Andrea, N. Gábor, K. János, N. Erick, Á. Bence, E. György, and G. György, “Water resources in efficient networks (wren) project overview and presentation of the research results carried out during the first year,” *AIS 2023 - 18th International Symposium on Applied Informatics and Related Areas*, pp. 90–98, 2023.

[4] P. Ramsey and V.-B. Columbia, “Introduction to postgis,” *Refractions Research Inc*, pp. 34–35, 2005.

[5] R. Obe and L. S. Hsu, *PostGIS in action*. Simon and Schuster, 2021.

[6] K. P. Gaffney, M. Prammer, L. Brasfield, D. R. Hipp, D. Kennedy, and J. M. Patel, “Sqlite: past, present, and future,” *Proceedings of the VLDB Endowment*, vol. 15, no. 12, 2022.

[7] J. Kreibich, *Using SQLite*. " O’Reilly Media, Inc.", 2010.

[8] M. Raasveldt, H. Muehleisen *et al.*, “Duckdb,” in *Proceedings of the 2019 International Conference on Management of Data. ACM*, 2019.

POSTER SESSION III.

Industrial and Embedded Systems

Possible Impacts of Ion Balance in Indoor Environment with Practical Application

Kristóf Koppány Enczi
Obuda University
Alba Regia Faculty
Székesfehérvár, Hungary
kristofenczi@stud.uni-obuda.hu

Bertalan Beszédes
Obuda University
Alba Regia Faculty
Székesfehérvár, Hungary
beszedes.bertalan@uni-obuda.hu
https://orcid.org/0000-0002-9350-1802

***Abstract*—This study focuses on ionizers, with particular emphasis on their effects on the human body and the environment, as well as the generation and composition of ions. Initially, the basic principles of ionizer operation will be explained, including the effects of negative ions on the human body and microorganisms. Following this, a detailed analysis of various ion generation methods will be provided, such as cosmic rays, UV radiation, natural and artificial corona discharge, lightning, the Lenard effect, and the ion-producing role of plants.**

The second part of the study concentrates on the specific design and construction of ionizers. The article thoroughly discuss the application of cascade multiplication and high-voltage rectification, high-voltage measurement methods, and the design of an "ion emitter." Subsequently, the implementation steps, including the development of circuit designs, the design process, the execution of measurements, the measurement of outgoing ions, and the experiences gained during the project will be presented.

Keywords – air ionizer, NAI, negative air ions, negative ion cluster

I. INTRODUCTION

The various sources of electromagnetic radiation, such as computers, televisions, mobile phones, and other electrical devices, have become an indispensable part of human life. These radiation sources can affect the human body in various ways, both directly and indirectly. The direct effect falls under the concept of "human exposure to EM fields" resulting from electromagnetic radiation sources. The human body reacts to exposure to electromagnetic fields, and this reaction depends on several factors, such as the type, intensity, and duration of the radiation. In contrast, electromagnetic radiation sources can also indirectly affect human health. For example, these devices can influence sleep patterns, stress levels, and contribute to changes in cognitive performance. In the past, when people suffered from lung diseases or other respiratory problems, the solution often involved breathing clean air, such as mountain air. Today, we observe and feel that urban air is significantly polluted, primarily due to exhaust gases, factory smoke, and other harmful substances in the air. This is important to mention because the aforementioned devices, especially in enclosed spaces, increase the generation of positive ions, and poor air quality further exacerbates this issue. An ionizer is a device that emits negative ions (anions) into the air in our environment, neutralizing the polluting particles found in the air of enclosed spaces. The device's function is to restore the poor air quality in our environment to a state similar to natural, fresh air. The use of this device is particularly recommended in homes, places with many children, clinics, hospital wards, offices, educational institutions, etc. In other words, places where many people spend long periods of time and where cations are prevalent.

II. THEROTICAL BACKGROUND

A. Air ions

Air ions are electrically charged molecules or atoms found in the atmosphere. An air ion is formed when a gas-phase molecule or atom gains enough energy to expel an electron. Negative air ions gain electrons, while positive air ions lose them. Natural and artificial energy sources include cosmic or radiative (ionizing) rays in the atmosphere; sunlight, including ultraviolet light; natural and artificial corona discharges, including lightning and thunder; the separation of water (Lenard effect); and energy sources from plants.

Radioactive elements like radium, actinium, and thorium emit α, β, and γ rays during their decay, ionizing the air. Cosmic radiation ionization is responsible for about 20% of ion generation in terrestrial areas [1]. Over the oceans, these sources generate negative air ions, with concentrations reaching up to 500 ions/cm³ on land and exceeding 1000 ions/cm³ at an altitude of 15 km [2].

The Earth's atmosphere is affected by a natural electric field, whose intensity varies under local and global influences [3]. Local effects include geographic location and weather conditions, while global factors refer to daily variations [1]. Sharp points on tree leaves or branches can cause corona discharge at high potential differences, emitting NAI. In mountainous areas, high electric fields and low air pressure promote corona discharge [4, 5]. Thunderstorms and lightning create very high electric fields, resulting in corona discharge, but NAI decreases once the thunderstorms subside. In forests, changes in the electric field during fog formation and dissipation can also induce corona discharge and NAI generation [3].

Artificial corona discharge is an effective method for generating NAI. Applying a high negative voltage to a conductor/electrode causes corona discharge if the electric field is strong enough [6]. The electric field is higher around the tip of a charged conductor with a sharp end, thus ionizing the air near the wire, creating negative air ions [7]. The intensity of the corona discharge depends on the shape, size of the wires, and the applied voltage. A sharp-pointed conductor results in more discharges, while larger diameter wires produce fewer; higher voltage generates more NAI [7]. NAI concentration is higher near the corona point, where an electron avalanche forms during corona discharge [8]. The application of electric fields and corona discharge to plants began in the 1960s [9]. Bachman and Hademenos (1971) demonstrated that high voltage resulted in more intense leaf tips, causing corona discharge and the formation of air ions and ozone [10]. The studies primarily focused on the biological effects, such as growth response, evaporation, plant damage, and the impact of produced ozone and negative air ions on growth [5].

B. *Composition of negative Air Ions*

Typically, negative air ions (NAIs) consist of several negatively charged molecules, and these negative ions can attach to multiple, or even 20-30 water molecules, forming negative ion clusters such as CO3−(H2O)n, O−(H2O)n, and O3−(H2O)n [11]. Mass spectrometry techniques have been widely used to determine the composition of NAIs from various sources [12]. Early measurements suggested that the majority of negative ions in the lower troposphere are O2−, CO3−, or NO3− and their (H2O)n clusters, as well as HSO4− core ions. Research has shown that atmospheric negative ions also contain additional ions such as OH−, NO2−, HCO3−, and their water clusters [13]. Fig. 1 and Fig. 2 summarizes the composition of NAIs from different sources and the evolution of oxygen-based NAIs.

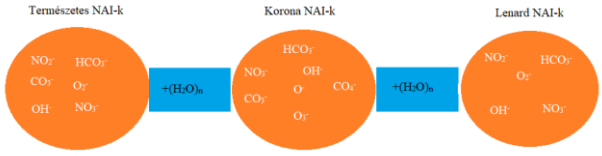


Figure 1: Composition of NAIs

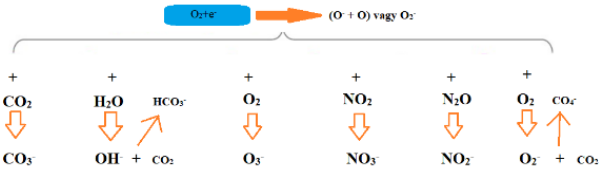


Figure 2: Development of oxygen-based negative air ions. Orange arrows indicate NAI transformation processes

C. *Mechanism of Action of Ionizers*

Ionizers used in homes produce negative and/or positive air ions (NAIs and PAIs), which can be generated simultaneously or separately. As a "side effect" of air purification, negative ions attract particles of dust, pollen, mold spores, tobacco smoke, and unpleasant odors in the rooms.

The ions produced by the ionizer are small and move at high speed, making them extremely effective. When these ions encounter pollutants, they transfer their charge to them. The resulting larger, charged particle is attracted to the nearest "grounded" surface, such as a wall or shelf where the ionizer is located. It is important for this surface to be washable, due to the "black wall effect."

As the ions move towards the interior of the room, their natural motion slows down. During this slowdown, the airborne pollutants (such as dust, pollen, cigarette smoke, and other gaseous pollutants) surround the ion, creating a "cluster." As a result, the ion's size increases significantly until it reaches a size that is too large and heavy to remain airborne, causing it to fall to the ground (Fig. 3).

The ionizer's ability to remove small particles from the air is particularly valuable for health. Recent research indicates that the smaller the pollutant particle size, the more difficult it is for the immune system to cope with it.

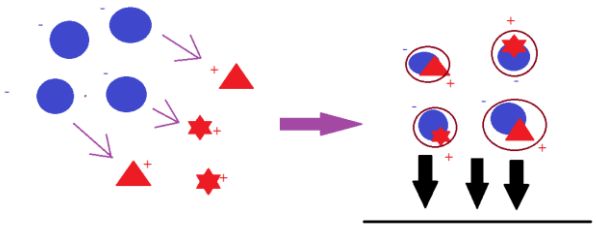


Figure 3: Size-dependent behavior of ions

D. *Ozone*

Ozone (O3) is a three-atom oxygen molecule that plays a significant role in the Earth's atmosphere. The ozone molecule consists of three oxygen atoms connected by strong bonds. It is a colorless gas but can appear bluish at high concentrations. It originates from both natural and human sources. Ozone was first discovered by its sharp smell and color; in 1840, Christian Friedrich Schönbein identified it while experimenting with oxygen and electrical sparks. The word "ozone" comes from the Greek "ozein," meaning to smell.

Ozone is present in different layers of the Earth. The ozone layer in the stratosphere absorbs the sun's harmful UV radiation, protecting terrestrial life from skin cancer and other health issues. In the atmosphere, ozone is found 10% in the troposphere (10-200 ppb) and 90% in the stratosphere (200-10000 ppb). The ozone shield absorbs a small portion of UV-B radiation and all UV-C radiation, enabling the sustenance of life.

Ozone is formed from oxygen under the influence of UV radiation, particularly UV rays with wavelengths shorter than 242 nm. Large amounts of ozone are produced around lightning strikes due to corona discharge. Tanning lamps filter out UV-C radiation with a special glass coating, as it also generates ozone. According to the World Health Organization, the maximum permissible concentration of ozone is 0.1 ppm [14].

Ozone is a highly oxidizing, toxic gas with a slightly unpleasant odor. Its byproducts are dangerous and carcinogenic, causing symptoms such as weakened immune system, reduced lung capacity, respiratory inflammations, and irritation. Ozone also acts as a greenhouse gas, preventing heat from returning to the sun. However, due to its disinfectant properties, it is beneficial as well, as its strong oxidizing ability effectively destroys viruses and bacteria. This property of O3 was discovered in 1886.

III. Implementation

A. *Circuit*

The method had been chosen involved a single pointed surface (formed from a pointed welding rod) connected to a potential ranging between -4 and -10 kV relative to ground. Due to the corona effect, the strong electric field that appears at the tip ionizes the air, as mentioned in previous chapters. It was necessary to ensure that the value did not exceed approximately -7 kV, because above this voltage, excessive corona discharge generates uncontrolled ozone, whose dangers have been previously discussed. It will be shown that the device did not exceed this value.

Among the solutions used here, the simplest possible one, which directly generates the desired voltage value from the

alternating current network using a voltage multiplier (see at Fig. 4.).

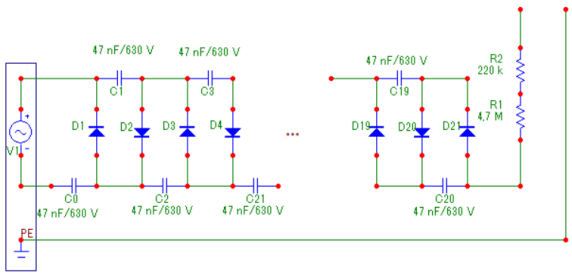


Figure 4: The schematic draw of the driver circuit

This construction is an 11-stage cascade multiplier (Fig. 5.), theoretically producing approximately -7 kV from the network (calculated value), and then stabilizing around -6 kV. Since higher voltages are generated at each stage, it was important for the capacitors to have a voltage tolerance of 630 V. The approximately 5 MΩ connected to the output was necessary to increase the internal resistance to ensure that any not intentional touch would be harmless. This approach is recommended for other circuits as well, and it is important that these resistors extend longitudinally and consist of more than just two parts, preferably several.

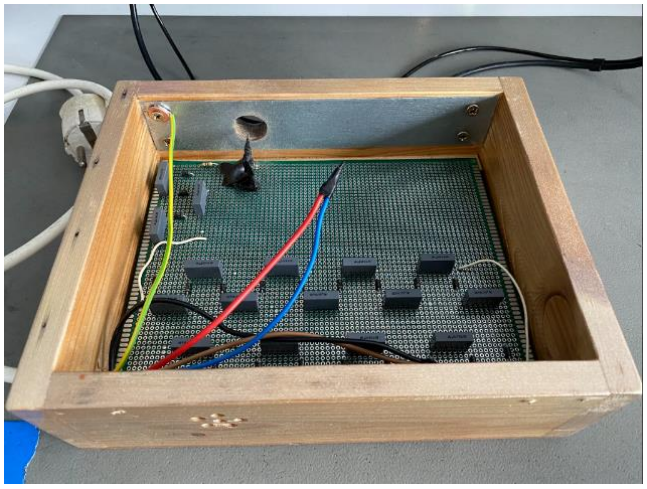


Figure 5: Physical appearance of the device

B. Measurement

The execution of the measurement required great attention due to the high voltage. Simple multimeters are not suitable for measuring high voltage, so a series resistor must be used to ensure the appropriate voltage for the instrument. An important characteristic of multimeters is the input resistance, which is dependent from the measurement range, and detailed in the datasheet.

To achieve a higher voltage range, the total resistance must be increased. In the case of high voltage, the internal resistance of the instrument is negligible, but it must be considered in lower ranges.Based on calculations, the appropriate shunt resistor was 2.7 MΩ, which achieved a 1:10 division. For the measurements, the Siglent SDM 3055 digital multimeter was available in the laboratory. The value measured at the output is shown in Fig. 6.

Figure 6: The measured value with shunt

Since it is on a 1:10 scale, the real value is multiplied by ten times, the number indicated by the device, so we get about -6kV (actually a negative voltage, which it was aiming for), perfectly matching expectations. This value should be higher, but the circuit has a high internal resistance, so the calculated value will differ from the measured value. Worth to note, that the described measurement method can only be a used for a well approximated indication.

C. Ion detector

To detect the emitted negative air ions, a simple circuit had been used (see on Fig. 7.). In this implementation, the leg of resistor R_3 serves as an antenna, which is sufficient for proper operation. The accumulation of ions on the antenna causes a small negative current, which flows to the base of transistor T_1.

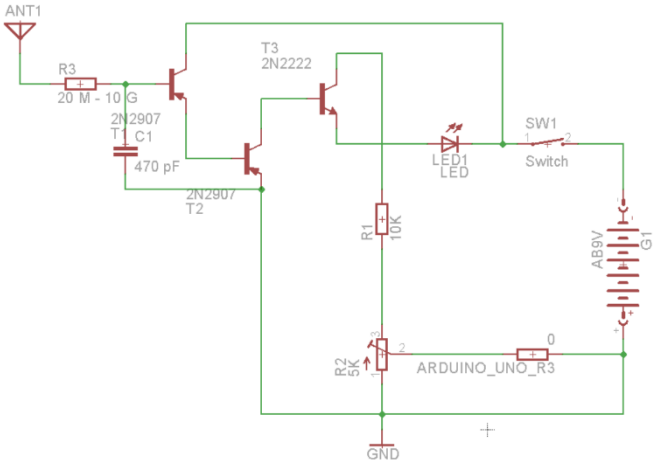


Figure 7: The scheamtic of ion detector

The capacitor C_1 and resistor R_3 function as an RC network, eliminating rapid fluctuations. If the negative current is large enough, it switches on T_1, which connects the negative pole of the B_1 battery (a TOE-8951 laboratory power supply during measurements) to the base of T_2, biasing and switching on T_2. This connects the base of T_3 to the positive pole of the battery, biasing and turning on T3. When T3 is switched on, the Arduino UNO R3, with the help of a shunt resistor and the internal ADC periphery, indicates the level of ion current, while LED lights up, indicating a strong ion field. Proper grounding is generally necessary for the unit to

function correctly, which will be connected to the ionizer's ground - in this case, a portable PSU is proposed.

After connecting the microcontroller and using the appropriate shunt resistor (in a similar way as in [15-18]), and then used the built-in serial plotter in the Arduino IDE (Fig. 8) to display the measured values in graph form. The previous measurements confirmed that it was possible to obtain adequate results, as first, the circuit current was measured with a digital multimeter. The measured range is in microamperes, so it had been selected the shunt resistor accordingly and wrote the code for the microcontroller. Similar control systems are detailed in [19, 20]. The applications of these control systems extend to advanced engineering fields [21, 22].

It is logical that the value changes as a function of distance; in this case, when the power of the ionizer was increased, the current also increased. This can be explained by the fact that there is a higher concentration of negative charges near the tip of the needle, resulting in a higher current flowing through the circuit. Fig. 9. shows an input power based measurement. As an opportunity for further development, it would be useful to collect spatial data, process it and graphically visualize the 3D environment for optimization [23, 24].

The ion detector can quickly signal the presence of a negative ion field, help identify its source, and indicate its relative strength, but it is not designed to measure absolute flux intensity. The ion detector can be used to filter out residual ion fields, check for ion leakage (e.g., during shielding tests), or test for static charges (in people's clothing, fluorescent tubes, plastic containers, certain winds, etc.), as well as for numerous other applications.

Figure 8: The measured values with Arduino

D. Current Measurement Instead of Voltage Measurement

The system is powered by the 230V public network. The microcontroller receives power from a 5V DC adapter. The power regulation of the ionizer is solved with a closed feedback loop, where the intervenor is a phase-splitting power regulator realized with a triac, galvanically isolated from the microcontroller by an optodiac, and the sensor is the circuit already presented (see the results in Fig. 10.). Comparable control loops and embedded systems are described in [25-27].

The output of the sensor is connected to the ADC of the microcontroller through the outputs of the terminals of a shunt. The current flowing through the shunt drops the voltage, which is measured by the microcontroller and displayed on the development interface of the serial plotter see on Fig. 11.

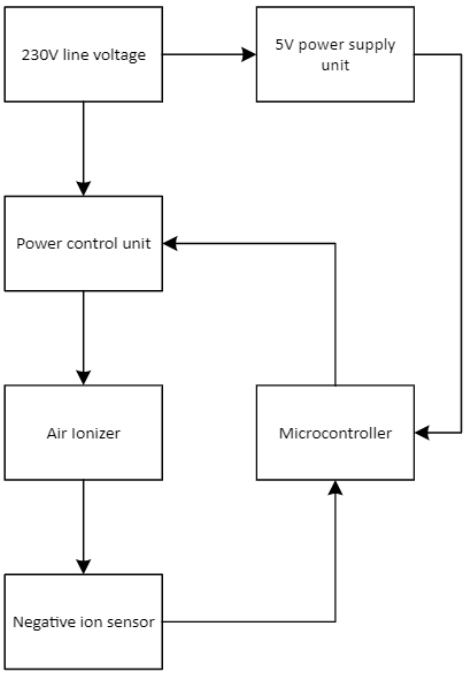


Figure 9: The block diagram of the project

Figure 10: Measured voltage, displayed in microamps

E. Measurement and Operational Experiences of the Ionizer

After measured the other stages as well and observed that the multiplication provided by the circuit is offset by the loss of each section. It had been measured roughly the same values at each stage. It is likely that the diy soldering is not optimal, and some voltage leaks from the system, resulting in reduced performance.

The device effectively collects positively charged particles, as shown in Fig. 11. This stage was made in a wooden box. The device works as expected, and the so-called ion breeze, mentioned in previous chapters, is noticeable. It produces a minimal amount of ozone, which is difficult to avoid, but plans include its detection and regulating the incoming voltage.

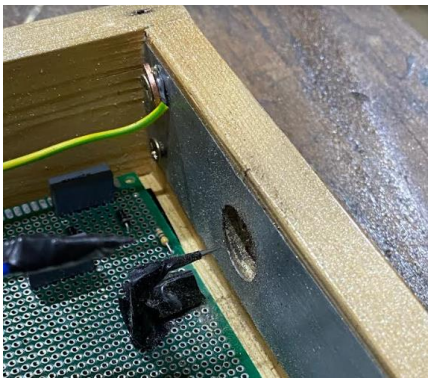


Figure 11: The result of the operation. As expected, the dust particles collect on the metal plate opposite the tip of the needle.

CONCLUSION

The study focused on the design, implementation, and effects of the ionizer, which plays a crucial role in air purification by generating negative and/or positive air ions. The ionizer, constructed using a simple yet effective 11-stage cascade multiplier, successfully generated the desired voltage of approximately -6 kV from the network. This was achieved despite the challenges posed by the high internal resistance of the circuit, which required careful measurement and calibration activities.

The main mechanism of the ionizer involves ionizing particles in the air through a high-voltage conductor. This process effectively attracts and neutralizes airborne contaminants, including dust, pollen, and smoke, thus purifying the air. Although the device produced a small amount of ozone, which is a known byproduct of ionization, the ozone levels were regulated and remained within safe limits. The presence of ozone highlights the need for further refinement in the design to minimize ozone production without compromising the efficiency of the ionizer.

Practical experiments and measurements confirmed that the ionizer effectively collects positively charged particles, reinforcing theoretical expectations. Despite some potential issues with soldering and voltage leakage in the circuit, the overall performance of the device was satisfactory.

In conclusion, the ionizer presents a promising solution for air purification, particularly in indoor environments. Its ability to remove fine particles from the air is beneficial not only for general cleanliness but also has significant health impacts, especially in areas with poor air quality. Future work should focus on optimizing the design of the ionizer to reduce ozone production and further improve the efficiency and safety of the device.

ACKNOWLEDGMENT

The authors would like to thank all the faculty staff and member that provide help and assistance throughout the project completion.

REFERENCES

[1] Hoppel W.A., Anderson R.V., Willet J.C. „Atmospheric Electricity in the Planetary Boundary Layer.” The Earth’s Electrical Environment. Washington, DC, USA: NAS Press, 1986. 195–205.

[2] Ermakov V.I., Bazilevskaya G.A., Potrevsky P.E., Stozhkov Y.I. „Ion balance equation in the atmosphere.” J. Geophys. Res. 102 (1997): 23413–23419.

[3] Borra J.P., Roos R.A., Renard D., Lazar H., Goldman A., Goldman M. „Electrical and chemical consequences of point discharges in a forest during a mist and a thunderstorm.” J. Phys. D: Appl. Phys. 30 (1997): 84–93.

[4] Harrison R.G., Carslaw K.S. „Ion-aerosol-cloud processes in the lower atmosphere.” Rev. Geophys. (2003): 41.

[5] Aubrecht L., Stanek Z., Koller J. „Corona discharge on coniferous trees—Spruce and pine.” Europhys. Lett. 53 (2001): 304–390.

[6] Altamimi G., Illias H.A., Mokhtar N., Mokhlis H., Bakar A.H.A. „Corona discharges under various types of electrodes.” Kuching Sarawak, Malájzia: Proceedings of the 2014 IEEE International Conference on Power and Energy (PECon), 2014.

[7] Berendt A., Budnarowska M., Mizeraczyk J. „DC negative corona discharge characteristics in air flowing transversely and longitudinally through a needle-plate electrode gap.” J. Electrostat. 92 (2018): 24-30.

[8] Manabe Y., Shimazaki T. „Formation mechanism of surface corona on dielectric plates under negative impulse voltage in atmospheric air.” IEEE Trans. Dielectr. Electr. Insul. 11 (2004): 631–637.

[9] L.E., Murr. „Plant growth response in a simulated electric field-environment. .” Nature 200 (1963): 490-491.

[10] Bachman C.H., Hademenos D.G. „Ozone and air ions accompanying biological applications of electric fields.” J. Atmos. Terr. Phys. 33 (1971): 497–505.

[11] Iwama H., Ohmizo H., Furuta S., Ohmori S., Watanabe K., Kaneko T., Tsutsumi K. „Inspired superoxide anions attenuate blood lactate concentrations in postoperative patients.” Crit. Care Med. 30 (2002): 1246–1249.

[12] J.G., Dillard. „Negative ion mass spectrometry.” Chem. Rev. 73 (1973): 589–643.

[13] Cabane M., Playe P. „Mass spectra of negative ions in air-like gas mixtures at atmospheric pressure.” J. Aerosol Sci. 11 (1980): 475–482.

[14] WHO, „National Library of Medicine,”National Center of Biotechnology Information, 2015. [Online 2023]. https://www.ncbi.nlm.nih.gov/books/NBK574591/#ch3.s4.

[15] S. Alsarayreh and Z. Sütő, "Power Converter Half-bridge Models in the Light of Real-time FPGA Implementation," *2021 IEEE 19th International Power Electronics and Motion Control Conference (PEMC)*, Gliwice, Poland, 2021, pp. 153-160, doi: 10.1109/PEMC48073.2021.9432622.

[16] Snezana Dragicevic, Aleksandar Peulic, Miroslav Bjekic, Radojka Krneta. Measurement and Simulation of Energy Use in a School Building. Acta Polytechnica Hungarica Vol. 10, No. 2, 2013. ISSN: 1785-8860

[17] Mosavi Amirhosein, Bertalan Beszedes, Imre Felde, Nadai Laszlo, Gorji Nima E. Electrical characterization of CIGS thin-film solar cells by two- and four-wire probe technique. MODERN PHYSICS LETTERS B 2020 p. 2050102 , 16 p. (2020)

[18] György Györök, Baklanov Alexander E, Bertalan Beszédes. Extension of Nodal Voltage Method with the Thermosensing. In: Orosz, Gábor Tamás (szerk.) AIS 2017 - 12th International Symposium on Applied Informatics and Related Areas organized in the frame of Hungarian Science Festival 2017 by Óbuda University : Proceedings. Székesfehérvár, Magyarország : Óbudai Egyetem (2017) 204 p. pp. 201-204. , 4 p.

[19] György Györök, András Dávid, Nikoletta Tolner, Bertalan Beszédes, Dániel Cseh. Supervision of the operation of digital circuits by Embedded Microcontroller. In: Orosz, Gábor (szerk.) AIS 2018 - 13th International Symposium on Applied Informatics and Related Areas. Székesfehérvár, Magyarország : Óbudai Egyetem, Alba Regia Műszaki Kar (2018) pp. 29-35. , 7 p.

[20] Györök György, Tihomir Trifonov, Alexander E. Baklanov, Bertalan Beszédes, Svetlana V. Grigoryeva, Aizhan Zhaparova. A Special Robust Solution for Battery Based Power Supply. In: Orosz, Gábor Tamás (szerk.) 11th International Symposium on Applied Informatics and Related Areas (AIS 2016). Székesfehérvár, Magyarország : Óbudai Egyetem (2016) pp. 32-35. , 4 p.

[21] Attila Sáfár, Bertalan Beszédes. Educational Aspects of a Modular Power Management System. In: Orosz, Gábor Tamás (szerk.) AIS 2019 : 14th International Symposium on Applied Informatics and Related Areas organized in the frame of Hungarian Science Festival 2019 by Óbuda University. Székesfehérvár, Magyarország : Óbudai Egyetem (2019) pp. 163-166. , 4 p.

[22] György Györök, Bertalan Beszédes. Artificial Education Process Environment for Embedded Systems. In: Orosz, Gábor Tamás (szerk.)

9th International Symposium on Applied Informatics and Related Areas - AIS2014. Székesfehérvár, Magyarország : Óbudai Egyetem (2014) pp. 37-42. , 6 p.

[23] A.Krasavin, M.Tolykbayeva, A.Kussaiyn-Murat, A.Kadyroldina, D.Alontseva. Surface restoration based on scanning data. 18th International Symposium on Applied Informatics and Related Areas, AIS, Óbuda University November 16, 2023 Székesfehérvár, Hungary, pp 68-72.

[24] A.Kadyroldina, A.Kussaiyn-Murat, A.Krasavin, D.Alontseva. A new laser stripe detection algorithm for 3D scanning applications. Международный симпозиум Applied Informatics and Related Areas, AIS, Óbuda University November 17, 2022 Székesfehérvár, Hungary. - pp 93-97.

[25] Alexander Baklanov, Svetlana Grigoryeva, György Györök. Control of LED Lighting Equipment with Robustness Elements, Acta Polytechnica Hungarica (1785-8860 1785-8860): 15 3 pp 105-119 (2016)

[26] G. Gyorok and M. Mako, "Configuration of EEG input-unit by electric circuit evolution," 2005 IEEE International Conference on Intelligent Engineering Systems, 2005. INES '05., Spain, 2005, pp. 255-258, doi: 10.1109/INES.2005.1555168.

[27] Aizhan Zhaparova, Dimitry Titov, Alexander Y. Baklanov, Gyorgy Gyorok. Study of the Effectiveness of Switching-on LED Illumination Devices and the Use of Low Voltage System in Lighting. Acta Polytechnica Hungarica (1785-8860 1785-8860): 12 5 pp 71-80 (2015)

Enhancing Analog Laboratory Power Supplies Functionality Using Microcontroller

Tóth Ágoston
Obuda University
Alba Regia Faculty
Székesfehérvár, Hungary
toth.agoston@stud.uni-obuda.hu

Bertalan Beszédes
Obuda University
Alba Regia Faculty
Székesfehérvár, Hungary
beszedes.bertalan@uni-obuda.hu
https://orcid.org/0000-0002-9350-1802

Abstract — This paper explores the enhancement of analog laboratory power supplies through the integration of microcontroller-based technology to upgrade and expand their user interfaces. Analog power supplies traditionally offer basic functionality with limited user interaction capabilities. This approach focuses on improving user accessibility and functionality by introducing a microcontroller-driven interface. The study concludes with a discussion on the potential applications and benefits of integrating microcontroller technology into analog laboratory power supplies, paving the way for more sophisticated and adaptable laboratory equipment in research and educational environments.

Keywords — analog power supply, embedded PSU, PSU UI, power supply user interface, PSU microcontroller

I. INTRODUCTION

The upgraded user interface enables precise control of output parameters such as voltage and current, offering users greater flexibility and accuracy in laboratory PSU settings. Additionally, advanced features such as digital display of output values [1-3], programmable presets [4-6], and remote control capability [7-9] are implemented to enhance usability and convenience.

Key aspects of the design include the selection of suitable microcontroller hardware [10-12] and development of intuitive firmware [13-16]. Practical implementation and experimental results demonstrate improvements in user experience and operational efficiency compared to conventional soulutions.

The goal of the development is to develop a device suitable for converting an analog power supply into a digital one. The goal is that the factory power supply should undergo as few modifications as possible. The subject is now an OMSZÖV OE-712 three-channel DC stabilized analog power supply. The device presented in the thesis can of course be used for any type of instrument. The solution is facilitated by the fact that the original documentation of the instrument is available, which describes its operation in detail, and also includes the circuit diagram and the PCB design.

II. DESCRIPTION OF THE TARGET INSTRUMENT

The OMSZÖV OE-712 instrument can be set to the desired value with the four (P1, P2, P3, P4) potentiometers on the front panel (see on Fig.1.). Replacing them with two two-channel digital potentiometers, the value of which will be set by the microcontroller. This solution involves the fewest changes within the instrument, because it is enough to replace only these limited number of parts.

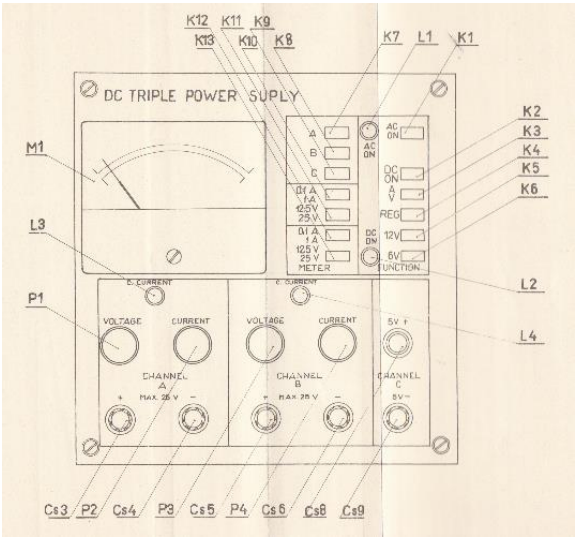


Figure 1: Drawing of the front panel of the laboratory power supply unit

In order to still be able to set the outputs manually, four rotary encoders are used. The digital potentiometers can be set based on their status. The principle of operation of the power supply unit (see on Fig. 2.) is that the filtered unstabilized direct voltage generated from the alternating voltage is regulated to the value of the set stable voltage by the pass-through transistor. The pass transistor is driven by operational amplifiers. The simplified diagram of the power supply is shown in Fig. 2. The mains voltage is connected to the mains transformer via a low-pass LC filter. The task of the input low-pass filter is to prevent high-frequency noise from the network to reach the output of the power supply. The mains transformer supplies the alternating voltage required for the device, after it is rectified, it also provides the voltage required for the output voltage and for powering the reference circuit and operational amplifiers.

Graetz rectifiers and capacitors produce a filtered, unstabilized direct voltage from the alternating voltages, from which the pass-through transistor produces a constant voltage and current with closed-loop control. The reference voltage required for comparison and generation of error signals is produced by a thermally compensated, high-stability IC. The heat-compensated Zener diode and

operational amplifier included in the two integrated circuit cases provide the reference voltage.

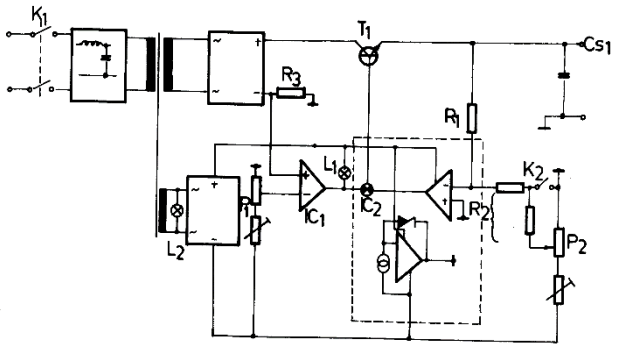


Figure 2: Block diagram of one channel of the laboratory power supply unit

III. REALIZATION

The goal is to make a PCB Shield (a circuit that can be easily attached to an open source board built around a microcontroller), which will contain the additional electronics and the connectors with which it is connected to the inside of the lab power supply and to the peripherals (LCD display, encoders, current measuring sensor). Thus ensuring the modular structure of the microcontroller base board, the additional Shield as a target PCB and the analog lab power supply system.

A. Digital potentiometer

The design started with the selection of digital potentiometers. From the technical documentation and circuit diagram of the laboratory power supply unit, it can be seen that the values of the voltage regulating potentiometers P7 and P15 are 4,7 KΩ, and the values of the current limit setting potentiometers P1 and P9 are 1 KΩ. The maximum output voltage is 25 V and the maximum output current is 1 A.

In addition to the above, the requirements imposed on the IC are the following. If the aim is to regulate the voltage with an accuracy of 0.1 V, it is necessary to be able to set the potentiometer in 25/0.1=250 steps, this requires the installation of an 8-bit resolution device (8 bits = 256 states). It is advisable to approximate the original 4,7KΩ values with the resistance values of the potentiometers.

The digital potentiometer should have a serial communication interface that is supported by the used microcontroller in terms of hardware and firmware, and should have THT housing to facilitate testability and development. Based on the conditions, the choice of components fell on the MCP4261 502E/P IC.

B. Current measurement

An ACS712 Hall cell current meter was installed for the current measurement. It operates on a 5V power supply, which is provided by a microcontroller panel with a 5V voltage regulator. As an output, it gives a voltage value, which is half of the supply voltage in the case of a current of 0 Ampere, and in the case of a change in the current strength of 1 Ampere, a voltage change of 0.185 V can be measured at the output. The output of the current meter can be read using the internal 12 bit analog/digital converter of the microcontroller. I_{max} = 5 Amps, which is suricient, even with a modified programmable ADC gain amplifier

The output of the power supply is fed back to one of the analog inputs of the microcontroller via a suitably sized voltage divider. The maximum output voltage of the lab power supply is 25 V, this must be distributed in such a way that the part that falls into the input of the microcontroller (in the case of Arduino Due, on board pins A10 and A11) is in no way greater than 3.3 V, since this is the connectable maximum voltage. In addition, the goal was to use high-value resistors so that the output would not be significantly loaded.

The output of the power supply is fed back to one of the analog inputs of the microcontroller via a suitably sized voltage divider. The maximum output voltage of the lab power supply is 25 V, this must be distributed in such a way that the part that falls into the input of the microcontroller (in the case of Arduino Due, on board pins A10 and A11) is in no way greater than 3.3 V, since this is the connectable maximum voltage. In addition, the goal was to use high-value resistors so that the output would not be significantly loaded.

25V/3.3V = 7.57 is the ratio of the voltage divider resistors, a high value available resistor is 150 KΩ, and 150 KΩ /7.57 = 19,82 KΩ ≈ 20 KΩ, which will be the other member of the resistor divider . In order for the measurement to be accurate, the exact values of the resistances must be known. Four-wire resistance measurements with a multimeter show the following: R9 = 19,6 KΩ, R10 = 143,6 KΩ, R11 = 18,8 KΩ, R12 = 136,2 KΩ.

The voltage of points A10 and A11 calculated from the above must be below 3.3 V at a voltage of 25 V, if this is not met, the values of the resistors must be modified. These results play an important role in writing the program and calibrating the system.

C. LCD display

The voltage and current values are displayed on a 2x16 character LCD display as shown in the figure below (Fig. 3.). Its management is simple, several program libraries are available for it. The LCD module and the microcontroller communicate via an I^2C communication bus. We connect the dedicated I^2C pins (20 SDA and 21 SCL) to the corresponding pins of the display. The 5V supply voltage required for operation is provided by the microcontroller development panel.

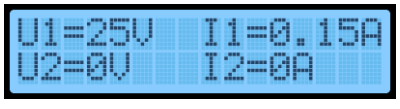


Figure 3: Extended user interface

D. System description

Fig. 4. shows the shield's final wiring diagram. For the sake of clarity, the externally placed connectors (pin rows) have a similar layout and have the same name as the socket to which they will be connected.

The most critical part of the PCB design is that the rows of pins with which the Shield is connected to the Arduino

board are in exactly the same position as the socket where they fit. To ensure this, the CAD files on the official Arduino website provide guidance when designing the board.

The rows of pins with which it is connected to the Arduino will face downwards (bottom side), and those from which the Shield is connected to the inside of the power supply will face upward (top side).

It is also possible to supply power to the Shield through the connector labeled POWER. This connects directly to the VIN and GND pins of the Arduino. Wider (0.5 mm) conductor strips were created for the power cables. The diagram below (Fig. 5) shows the PCB plan, the upper copper foil is shown in red and the lower copper foil is shown in blue.

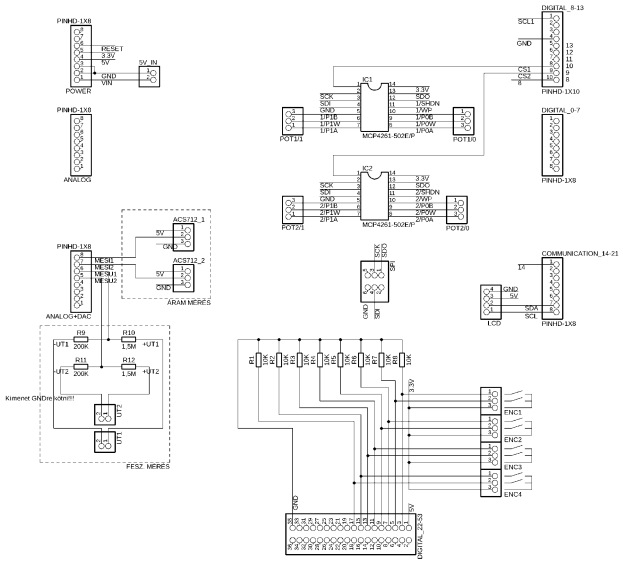


Figure 4: Circuit diagram of the auxiliary circuit

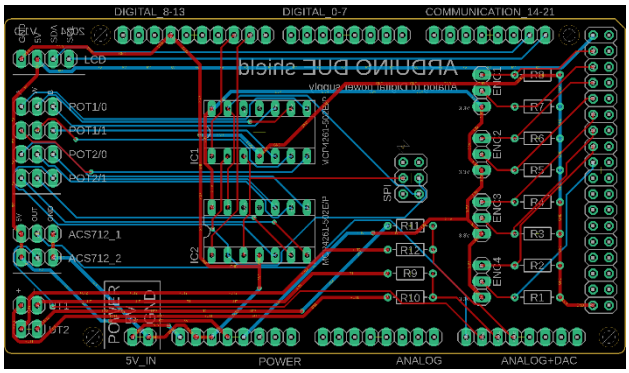


Figure 5: PCB plan of the auxiliary circuit

E. Shield és a tápegység összekötése

Instead of the factory potentiometers, the outputs of the digital potentiometers must be connected (POT1/0, POT1/1, POT2/0, POT2/1 connectors on the shield, see on Fig. 6.), which perform the following functions:

- POT1/0: CHA voltage adjustment,
- POT1/1: CHA current adjustment,
- POT2/0: CHB voltage adjustment,
- POT2/1: CHB current adjustment.

One current sensor had been connected in series with the output of channels A and B. They must be connected to the corresponding connector on the Shield (ACS712_1 and ACS712_2). The UT1 and UT2 connections are connected to channels A and B of the power supply in parallel.

The PCB was made in a PCB factory based on the plans, after it arrived it just had to be assembled and attached to the Arduino board, on which it fit perfectly (Fig. 7).

The wires can be distinguished from each other based on their color. A pinheader has been soldered to the end of each wire bundle, making it easy to connect them. Inside the power supply unit, the custom-made wire harness had to be soldered to the factory internal wires.

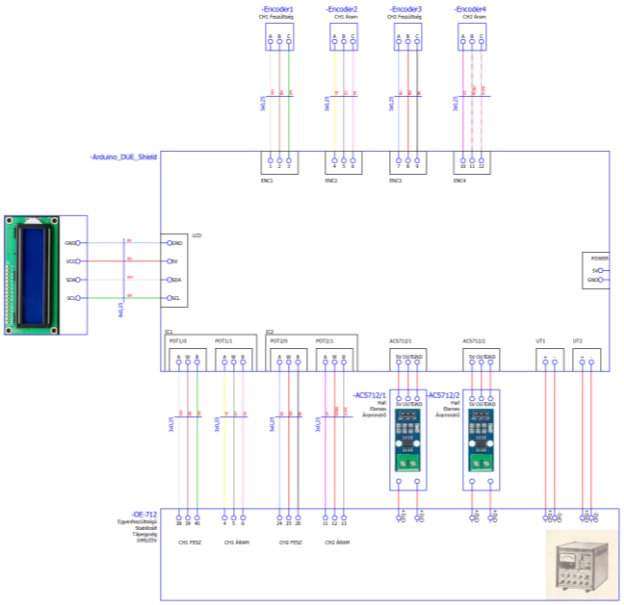


Figure 6: System architecture

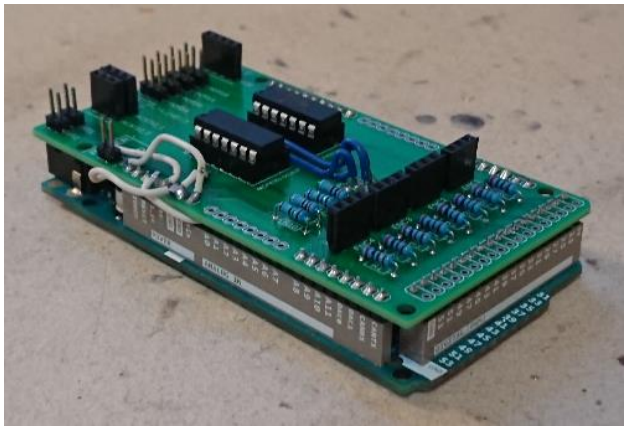


Figure 7: The developed shield

IV. FIRMWARE IMPLEMENTATION

A. Calibration of the current meter

The most important parts of the routines uploaded to the Arduino are presented below. The function performs the calibration of the current sensors (offset compensation operation) (Fig. 8). During the calibration, the program performs 50 samples on the two specified analog inputs (channel1 and channel2), arranges them in a row, discards the upper and the lower 10-10 values, and then calculates

the average of the remaining 30 values (it is important that before calling it to 0 the output current of the power supply must be adjusted). It subtracts the value of the calculated average from the value 2048 (12-bit samples are taken → 4096/2 = 2048) and uses this value as an offset. The value of the current sensor corresponding to 0 amperes is calculated with the zeroAmpere variable, adding the offset.

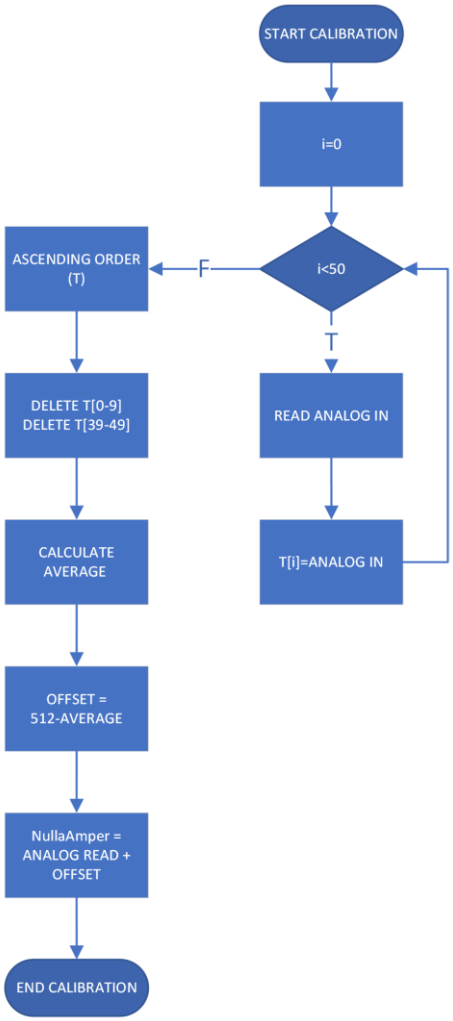


Figure 8: Current sensor calibration

B. *MCP4261 setting functions*

These functions describe the commands that control the MCP4261 digital potentiometers using SPI communication. The desired IC is selected with the CS (Chip Select) pin, and then the desired command to increase or decrease the resistance values is sent using the SPI.transfer function. First, it is necessary to pull the Select input of the IC to be set to logic low, then issue the increment or decrement command.

C. *void loop()*

The setup function runs once after switching on. The outputs and inputs are set here, followed by the main program.

Zero voltage1 and zero voltage2 calculate the zero point of the current sensors. The output value of the current sensors is converted into voltage values by compensating the offset and multiplying by the appropriate scale factor (using a reference voltage of 3.3 V). This voltage is saved in the zerovoltage1 and zerovoltage2 variables.

ActualVoltage1 and actualVoltage2 calculate the current voltage from the current sensors (Fig. 9). For this, the output value of the current sensor is added to the offset-compensated value, and then also multiplied by the scale factor, it is converted into a voltage value.

Current1 and current2 calculate the current value from the current sensors (Fig. 9). The current voltage subtracted from the zero point is divided by the current constant (0.185 A/V) to calculate the current.

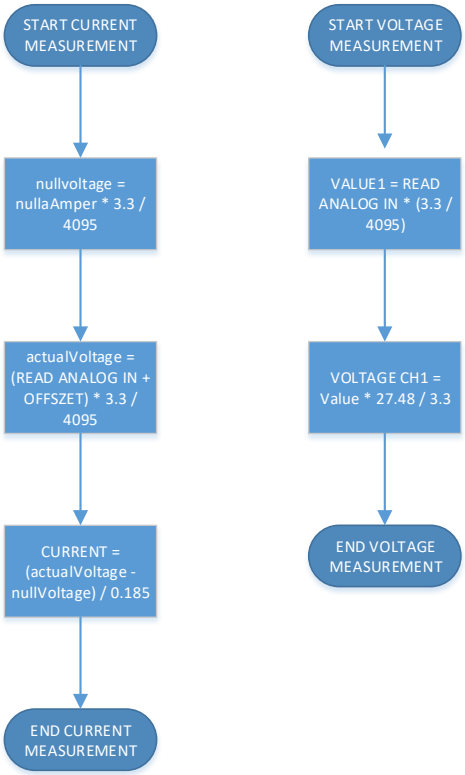


Figure 9: Current and voltage measure

This part is responsible for calculating voltage values. In the code fragment, it reads the values from two analog inputs (A10 and A11) and then represents them as voltage values.

Value1 and value2 normalize the values read from the analog inputs with the factor (3.3/4095.0) to obtain the voltage values from the analog values measured in the range 0-4095. CH_Voltage will be the currently measured voltage.

After this, there is the function that writes out the data, which will not be detailed now. After that, the program monitors the different states of the encoders (outputA and outputB) and then decides to increase or decrease the state of the potentiometers based on these (Fig. 10). To increase or decrease the potentiometers, use the functions increment0, increment1, decrement0, decrement1, which control the MCP4261 IC.

The implemented prototype circuit, shown in Fig. 11., is a microcontroller-driven system developed and successfully deployed during the research. This system

precisely controls the operation of the power supply, allowing for fine-tuning of output voltage and current through digital adjustments. The analog components integrated within the circuit provide a stable, low-noise output, while the microcontroller control offers flexibility and programmability, enabling the user to select various setting profiles and automatic voltage regulation options. The device is suitable for laboratory measurement and testing tasks, where precise voltage and current values are critical, and it is particularly valuable for educational purposes, as it effectively demonstrates the operating principles of programmable power supplies and the advantages of analog-to-digital interface applications.

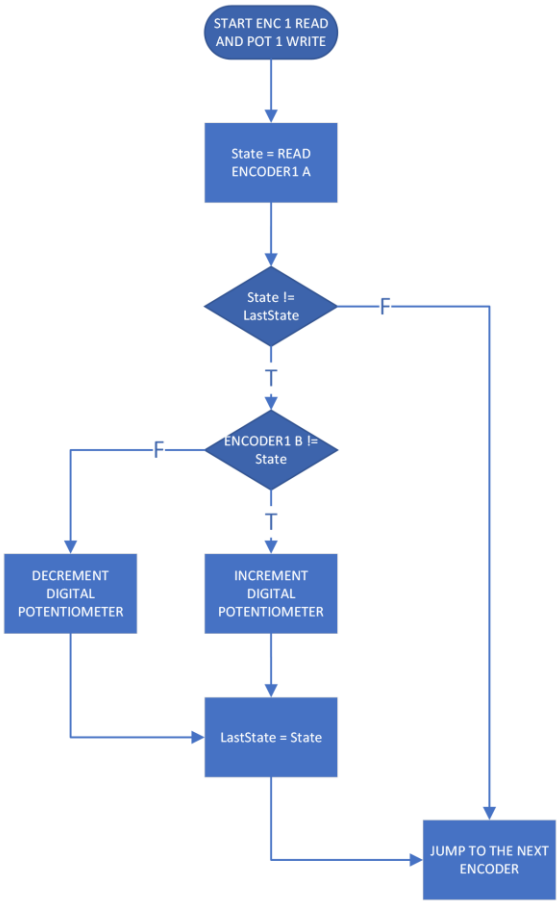


Figure 10: Encoders and digital potentiometers setup function

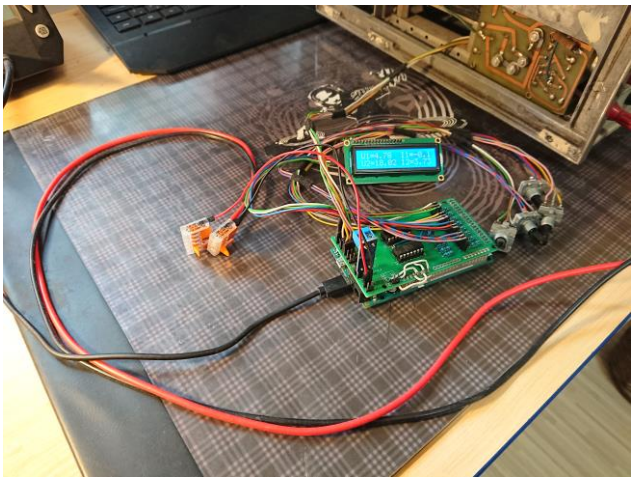


Figure 11: The assembled prototype of the device

Conclusion

In conclusion, the integration of microcontroller technology into analog laboratory power supplies significantly enhances their user interfaces and overall functionality. By upgrading the user interface, it is introduced precise control over output electrical parameters, digital displays, programmable presets, and remote control capabilities, thereby transforming traditional analog power supplies into more versatile and user-friendly devices. The practical implementation and experimental results confirm that these enhancements lead to improved user experience and operational efficiency. The microcontroller-based upgrades provide greater flexibility, accuracy, and convenience, making the power supplies more suitable for a variety of applications in research and educational settings.

Acknowledgment

The authors would like to thank all the faculty staff and member that provide help and assistance throughout the project completion.

References

[1] S. Alsarayreh and Z. Sütő, "Power Converter Half-bridge Models in the Light of Real-time FPGA Implementation," *2021 IEEE 19th International Power Electronics and Motion Control Conference (PEMC)*, Gliwice, Poland, 2021, pp. 153-160, doi: 10.1109/PEMC48073.2021.9432622.

[2] Snezana Dragicevic, Aleksandar Peulic, Miroslav Bjekic, Radojka Krneta. Measurement and Simulation of Energy Use in a School Building. Acta Polytechnica Hungarica Vol. 10, No. 2, 2013. ISSN: 1785-8860

[3] Mosavi Amirhosein, Bertalan Beszedes, Imre Felde, Nadai Laszlo, Gorji Nima E. Electrical characterization of CIGS thin-film solar cells by two- and four-wire probe technique. MODERN PHYSICS LETTERS B 2020 p. 2050102 , 16 p. (2020)

[4] György Györök, Baklanov Alexander E, Bertalan Beszédes. Extension of Nodal Voltage Method with the Thermosensing. In: Orosz, Gábor Tamás (szerk.) AIS 2017 - 12th International Symposium on Applied Informatics and Related Areas organized in the frame of Hungarian Science Festival 2017 by Óbuda University : Proceedings. Székesfehérvár, Magyarország : Óbudai Egyetem (2017) 204 p. pp. 201-204. , 4 p.

[5] György Györök, András Dávid, Nikoletta Tolner, Bertalan Beszédes, Dániel Cseh. Supervision of the operation of digital circuits by Embedded Microcontroller. In: Orosz, Gábor (szerk.) AIS 2018 - 13th International Symposium on Applied Informatics and Related Areas. Székesfehérvár, Magyarország : Óbudai Egyetem, Alba Regia Műszaki Kar (2018) pp. 29-35. , 7 p.

[6] Györök György, Tihomir Trifonov, Alexander E. Baklanov, Bertalan Beszédes, Svetlana V. Grigoryeva, Aizhan Zhaparova. A Special Robust Solution for Battery Based Power Supply. In: Orosz, Gábor Tamás (szerk.) 11th International Symposium on Applied Informatics and Related Areas (AIS 2016). Székesfehérvár, Magyarország : Óbudai Egyetem (2016) pp. 32-35. , 4 p.

[7] Attila Sáfár, Bertalan Beszédes. Educational Aspects of a Modular Power Management System. In: Orosz, Gábor Tamás (szerk.) AIS 2019 : 14th International Symposium on Applied Informatics and Related Areas organized in the frame of Hungarian Science Festival 2019 by Óbuda University. Székesfehérvár, Magyarország : Óbudai Egyetem (2019) pp. 163-166. , 4 p.

[8] György Györök, Bertalan Beszédes. Artificial Education Process Environment for Embedded Systems. In: Orosz, Gábor Tamás (szerk.) 9th International Symposium on Applied Informatics and Related Areas - AIS2014. Székesfehérvár, Magyarország : Óbudai Egyetem (2014) pp. 37-42. , 6 p.

[9] I. Bendiák and T. Sándor, "Possible Ways of Measuring and Calculating Waste Heat from a Machine Diagnostic Approach," *2023 IEEE 6th International Conference and Workshop Óbuda on Electrical and Power Engineering (CANDO-EPE)*, Budapest,

Hungary, 2023, pp. 000259-000266, doi: 10.1109/CANDO-EPE60507.2023.10417972.

[10] I. Bendiák and S. Semperger, "Simplified Predictive Strategy of Mechanical Life Cycle Model in Three-Phase Asynchronous Motor," *2023 IEEE 6th International Conference and Workshop Óbuda on Electrical and Power Engineering (CANDO-EPE)*, Budapest, Hungary, 2023, pp. 000267-000274, doi: 10.1109/CANDO-EPE60507.2023.10417977.

[11] Alexander Baklanov, Svetlana Grigoryeva, György Györök. Control of LED Lighting Equipment with Robustness Elements, Acta Polytechnica Hungarica (1785-8860 1785-8860): 15 3 pp 105-119 (2016)

[12] G. Gyorok and M. Mako, "Configuration of EEG input-unit by electric circuit evolution," *2005 IEEE International Conference on Intelligent Engineering Systems, 2005. INES '05.*, Spain, 2005, pp. 255-258, doi: 10.1109/INES.2005.1555168.

[13] Aizhan Zhaparova, Dimitry Titov, Alexander Y. Baklanov, *Gyorgy* Gyorok. Study of the Effectiveness of Switching-on LED Illumination Devices and the Use of Low Voltage System in Lighting. Acta Polytechnica Hungarica (1785-8860 1785-8860): 12 5 pp 71-80 (2015)

[14] S. Kimathi, B. Lantos. Modelling and Attitude Control of an Agile Fixed Wing UAV based on Nonlinear Dynamic Inversion, Periodica Polytechnica Electrical Engineering and Computer Science, Vol 66(3), 2022, pp. 227-235

[15] S. Kimathi, B. Lantos. Quaternion based attitude control of a maneuvering fixed wing UAV, IEEE 26th International Conference on Intelligent Engineering Systems (INES), pp. 261-266, 2022.

[16] Thermo-dynamic Cycle Computation of a Micro Turbojet Engine. Fozo, L; Andoga, R and Kovacs, R. 17th IEEE International Symposium on Computational Intelligence and Informatics (CINTI). 2016 17TH IEEE International Symposium on Computational Intelligence and Informatics (CINTI 2016) , pp.75-79

Fuzzy Logic Controller for Regulating the Indoor Temperature

Tímea Bekk
Alba Regia Faculty
Obuda University
Székesfehérvár, Hungary
bekk.timea@amk.uni-obuda.hu

Éva Takács
Alba Regia Faculty
Obuda University
Székesfehérvár, Hungary
takacs.eva@amk.uni-obuda.hu

Péter Tarsoly
Alba Regia Faculty
Obuda University
Székesfehérvár, Hungary
tarsoly.peter@amk.uni-obuda.hu

Rozália Lakner
Alba Regia Faculty
Obuda University
Székesfehérvár, Hungary
lakner.rozalia@amk.uni-obuda.hu

***Abstract*— In the article, we present the implementation of indoor temperature control with a fuzzy controller using Matlab Fuzzy Logic Toolbox. In order to validate and test the proposed method, a MATLAB-Simulink model of the system under study has been prepared, in which the fuzzy controller has been integrated. The resulting fuzzy controller was compared with the most used controllers and the results were evaluated.**

Keywords – Fuzzy logic controller, Heating system regulation, Matlab Fuzzy Logic Toolbox, Matlab Simulink

I. INTRODUCTION

In the article, the different implementation options for temperature control of a room are presented, especially the fuzzy controller that works with linguistic expressions. We examined a specific room, located at 1-3 Székesfehérvár Pirosalma street 103 room on the 1st floor of the building, we used its parameters in my models. In the Simulink program, we will create a model for heating control, in which we will create three types of control methods. These are: ON/OFF controller, PID controller and Fuzzy controller. When creating the model, we will take into account the physical size of the room, including the surface of the walls and windows. We examine the possibility of setting the desired temperature (set point) and the limit of outside temperature fluctuations.

We will examine the results of the simulation of the three implemented control systems and evaluate them by comparing them.

II. FUZZY LOGIC CONTROLLER

A. *Fuzzy systems*

Fuzzy logic systems use calculation methods that not only use traditional true/false logic, but also handle linguistic variables and the uncertainty inherent in them. Fuzzy systems are built from fuzzy sets, the elements of which can take not only the values 0 or 1, but also any intermediate value to specify the strength of belonging to the set.

One of the main goals of introducing fuzzy logic was to better model uncertain concepts, natural language concepts, and human thinking. Fuzzy sets characterized by non-sharp boundaries were described by Zadeh in an article published in 1965 [1], and in 1973 he created the first fuzzy inference algorithm using linguistic variables. This was still very complicated and required a lot of calculations, however, in 1975, Mamdani reformulated the language variables and thanks to the new solution, the calculation requirements were drastically reduced. [2]

B. *Fuzzy sets and their basic types*

We often come across expressions that cannot be clearly determined as true or false, elements or non-elements of a given set. If, for example, we want to describe the set of "pleasant room temperature", we should define a sharp set according to the classical set theory, for example by specifying a minimum and a maximum value, by which temperature values falling within a defined interval are included in the set, temperatures outside the interval values are not. Here, on the one hand, there may be a question as to what these awarded values should be. If an agreement was reached at all, then a further question may arise, that a temperature value that is, for example, 0.1 lower than the specified minimum value, or 0.1 higher than the specified maximum value, no longer belongs to the set of "pleasant room temperature". Fuzzy sets can be used to mathematically describe these types of expressions, in which the membership can be described by a membership function. The difference between a fuzzy set described by a membership function and a sharp set in classical set theory is illustrated in Fig. 1 by a concrete example, defining "pleasant room temperature" sets in the sense of temperature universal.

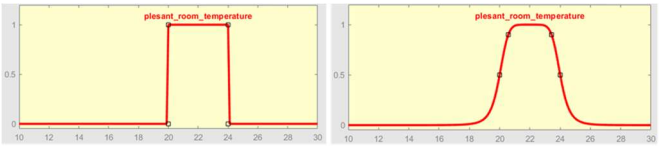


Fig. 1: Crisp set and fuzzy set.

C. *Fuzzy rules*

Fuzzy sets can be used to construct fuzzy rules, which can be formulated as conditional statements of the form “IF x = A AND y = B”, where A and B are fuzzy sets and x and y are the values of the corresponding universes.

An "ordinary" rule is fireable if the condition part is true and the firing results in the consequence part being enforced (its argument will be true).

A fuzzy rule will only fire to a certain extent, i.e. partially, and therefore the consequence part will only be true to a certain extent. This is because, for a fuzzy rule, the condition

part will be true with a value between 0 and 1, determined by the current input value. This is called the degree of fit. The firing of a rule requires that the measure of fit and the intersection of the fuzzy set in the condition part are not empty, and the consequence value is the combined measure of the measure of fit of the fuzzy set in the consequence part.

The rules can also provide expert knowledge and practical experience for a particular task. The set of rules for a given task forms a rule base, which can be used to determine the behavior of the system using fuzzy inference. The advantage of such a fuzzy model is that it can be described in linguistic terms, like human thinking, even without any knowledge of systems theory or control theory.

D. Fuzzy expert system

A fuzzy expert system is an expert system whose knowledge base is made up of fuzzy rules interpreted on fuzzy sets, and the inference machine maps fuzzy inputs into fuzzy outputs using fuzzy rules, in which fuzzy conclusions are logical processes based on the basic principles of fuzzy logic, as well as enabling they do the modeling of non-linear, in addition uncertain relationships.

If this system is operated with real inputs and outputs, then it must be supplemented with fuzzifier units that transform the real (crisp) input into a fuzzy input, and defuzzifier units that transform the fuzzy output into a real (crisp) output.

III. ROOM TEMPERATURE CONTROL WITH FUZZY CONTROLLER

A. The investigated room

To examine the temperature control task of our article, we chose one of the university's classrooms. The selected classroom is a relatively large lecture hall, which can accommodate approximately 60 students.

The main parameters of the room are as follows: floor area of the hall: 112.56 m^2, interior height: 3.20 m, internal volume: 360.19 m^3, number of windows: 5, type of radiators: plate radiator, number of radiators: 5 pcs, size of radiators: 1.60 m * 0.60 m * 0.09 m, orientation of the room: South, Southeast.

B. The investigated model

When examining the room subsystem, we are interested in the temporal behavior of the room temperature. Thus, when modelling the room, we consider the indoor temperature as a state variable, the change of which is influenced by the thermal energy transferred during heating on the one hand and the energy loss to the environment on the other hand. This can be expressed by the following equation of state:

$$\frac{dT_{room}V\rho c_p}{dt} = Q_{heati} \quad - Q_{loss}$$

where

T_{room}: temperature of the room [°C]

V: internal volume of the room [m^3]

ρ: density of air in the room [kg/m^3]

c_p: specific heat capacity of the air in the room [J/kg°C]

$Q_{heating}$: rate of heat energy transferred [J/s]

Q_{loss}: rate of energy loss [J/s]

Assuming a constant volume, density and specific heat capacity, the equation can be transformed as follows:

$$\frac{dT_{room}}{dt} = \frac{Q_{heating}}{V\rho c_p} - \frac{Q_{loss}}{V\rho c_p}$$

The amount of heat energy transferred depends on the difference between the temperature of the heater and the room

$$Q_{heating} = Q(T_{heater} - T_{room})$$

and the energy loss depends on the difference between the room temperature and the outside temperature

$$Q_{loss} = U_w A_w (T_{room} - T_{out})$$

where

T_{heater}: heater temperature [°C]

Q: heat rate [W/°C]

U_w: heat transfer coefficient [W/m^2 °C]

A_w: heat transfer surface [m^2]

T_{out}: the outside temperature [°C]

The value of the heat transfer coefficient was determined by taking into account the area of the wall and window on the front wall of the room and assuming average insulation.

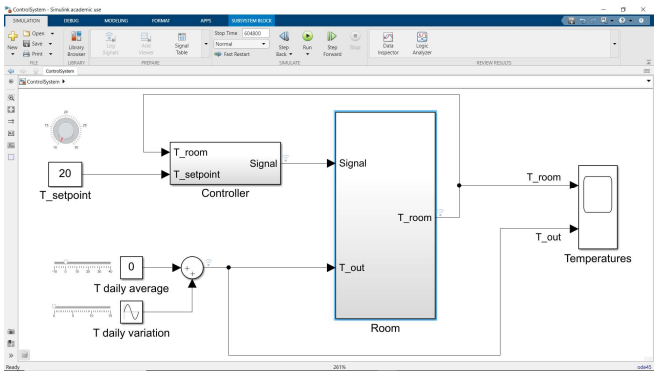


Fig. 2: Simulink model of the investigated system

The resulting equations were used to build the Matlab Simulink [3] model of the room subsystem (Fig. 2). Two input ports (the controller output, "Signal" and the outside temperature, "T_out") and one output port (the temperature of the room, "T_room") were designed for the subsystem. The subsystem and its elements are shown in Fig. 3.

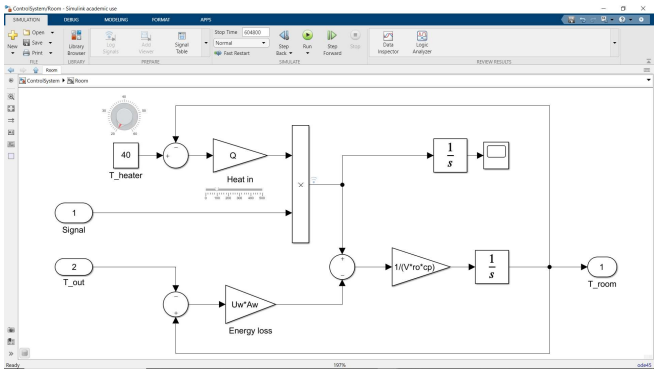


Fig. 3: Room subsystem and its blocks

The value of the "T_heater" block in the upper left corner of the figure shows the temperature of the heating device

(radiator). A constant value was assumed during the simulation, but by changing the value, it is also possible to examine the effect of different radiator temperature settings. The value set in the figure is 40 °C, which can be changed either manually using the "Knob" block above the block, or directly by changing the constant parameter of the "T_heater" block. For setting with the "Knob" block, values between 20 and 60 °C have been defined. (Fig. 3)

The next larger unit is the "Gain" block called "Heat in", which can be used to determine the value of the thermal energy transferred from the heating device to the room, which is proportional to the temperature difference between the heating device and the room. Under the "Heat in" block there is a scale on which the value of Q can be set manually. The value of the transferred thermal energy is also influenced by the controller's output signal ("Signal").

In the lower left part of the figure, you can see the definition of energy loss ("Gain" block named "Energy loss"), which represents the thermal energy leaving through the facade wall and windows of the examined room. This loss is proportional to the difference between the room temperature ("T_room") and the outside temperature ("T_out"). It is also influenced by the heat transfer surface ("Aw"), which in our case is 45.02 m^2, and the heat transfer factor ("Uw"). On the front wall of the room, 1/3 is wall, and 2/3 is window, taking this into account and assuming average insulation, the value of the latter is 0.94 W/m^2 °C. The difference between the heat energy transferred by heating and the heat energy loss gives the value of the heat energy change. From this, the temperature change can be calculated with another "Gain" block, whose amplification factor is 1/(V*ρ*cp), where the internal volume of the room is 360.19 m^3, ρ is the air density 1.205 kg/m^3, cp is the specific heat of the air 1013 J/ kg °C.

The 1/s "Integrator" block, which outputs the value of the time integral of the input signal, in this case the output is the temperature of the room ("T_room"), is used to check how much the temperature increases in a unit of time.

In the lower left corner of the Simulink model (Fig. 2), there is a subsystem created for the simulation of the outside temperature. It consists of two blocks. One can be used to set the daily average temperature ("T daily average"), which is set to 0 °C in the figure. The second block describes the daily fluctuation using a sine function. The value of the fluctuation in the case of the figure is 5 °C. These are constant values; however, they can be adjusted using the "Slider" blocks placed next to them, or by clicking on the "Constant" blocks to set the values.

To display the simulation, we placed a "Scope" block with the name "Temperatures" in the Simulink model (Fig. 2). This is a block that allows graphical display and monitoring of signals during testing. The "Scope" block provides the possibility to configure the display characteristics, we can also set the number of displayed signals, the time range and the color of the signals. Since we can display several signals at the same time, we have the opportunity to easily perform the comparison.

The graph displayed by the "Scope" block shows the time on the horizontal axis and the temperature on the vertical axis. During the simulation, it is possible to change the characteristics so that the change can be monitored. The input signals for the "Scope" block of the model are the room temperature ("T_room") and the outside temperature ("T_out").

In the upper left part of the Simulink model, there is the control subsystem and one of its inputs, the setpoint. The block called "T_setpoint" represents the expected room temperature, which was set to 20 °C in the case shown in Fig. 2. This can be changed in two ways. With the "Knob" block located above the setpoint, the interval value of which is defined between 10-30 °C. Here we can set the setpoint manually, too.

The "Controller" subsystem determines the output of the controller ("Signal") based on the input values ("T_room" and "T_signal"). As a controller subsystem, a fuzzy controller system was tested, whose operation was compared with two other temperature controllers.

C. *Fuzzy controller*

During the implementation of the fuzzy controller, the first step was to design the fuzzy system [5, 6], the corresponding inputs and the output. It is important what kind of membership functions we create so that it works well with our Simulink model later on. Therefore, we chose input variables whose value can be determined from the model.

Two inputs were created with the help of Matlab Fuzzy Logic Toolbox [4]. One is called "T_difference". This shows how much the room differs from the set temperature. For this, we defined four membership functions. (Fig. 4.)

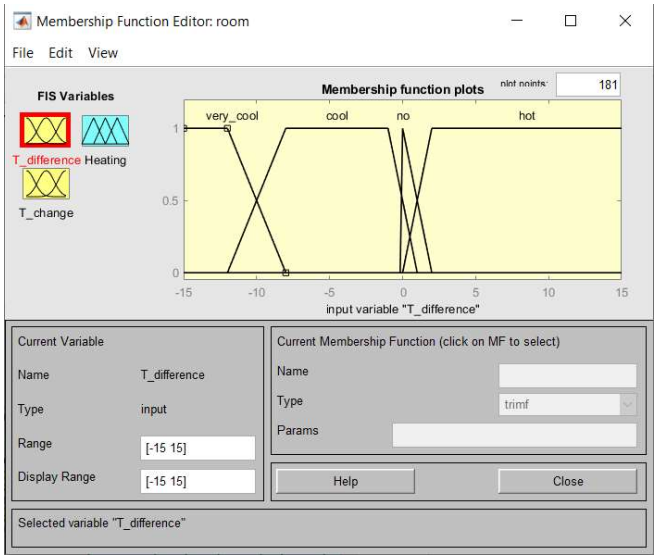


Fig. 4: Fuzzy sets of "T_difference".

We have made sure that there is overlap between the membership functions. In addition to the trapezoidal membership functions, the "no" membership function was given a narrow triangular shape because there is a narrow interval where the deviation is minimal, so we will not need to change the heating.

We gave the second input the name "T_change". This shows how the temperature of the room changes over a unit of time. We created 5 membership functions, four trapezoidal and one triangular curve. (Fig. 5.)

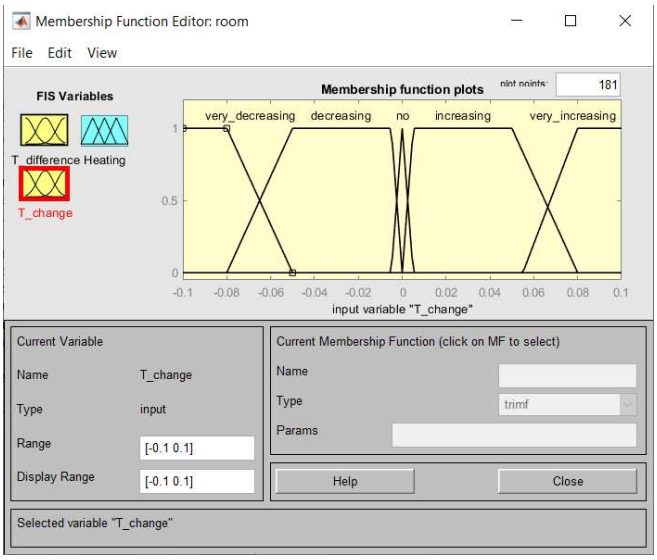


Fig. 5: Fuzzy sets of "T_change".

After that, we created the output. There is an output named "Heating". We created five membership functions, which are also trapezoidal and triangular curves. (Fig. 6.)

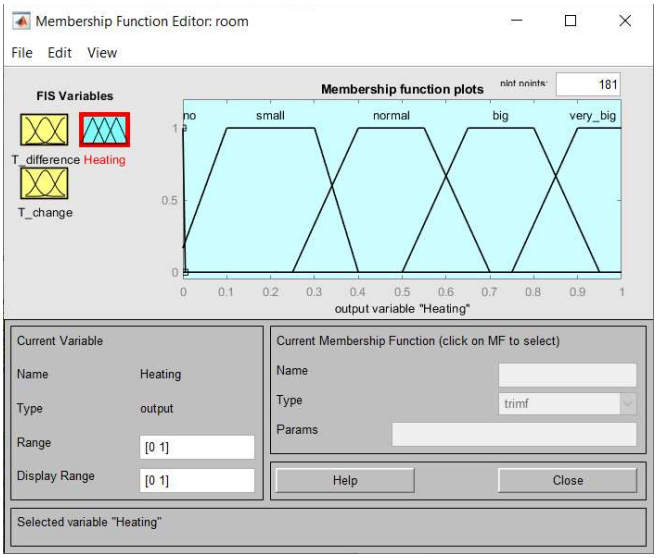


Fig. 6: Fuzzy sets of "Heating".

After designing and creating the input and output membership functions, the rules were created. We have established 12 rules, taking care to take into account all possibilities that may occur during temperature control.

The defined rules are shown in Fig. 7. using the View/Rules view of the Fuzzy Logic Toolbox. Each row represents a rule, so 12 rows are visible. The first two columns show the "if" part of the rule, and the last column shows the "then" part, the conclusion. The fuzzy combination takes place in the lowest table of the last column, where the result of the defuzzification is also displayed, i.e. the real value of the fuzzy control signal (with a red vertical line). It can be seen that if "T_difference" is equal to -5 and "T_change" is zero, then the value of "Heating" is 0.725.

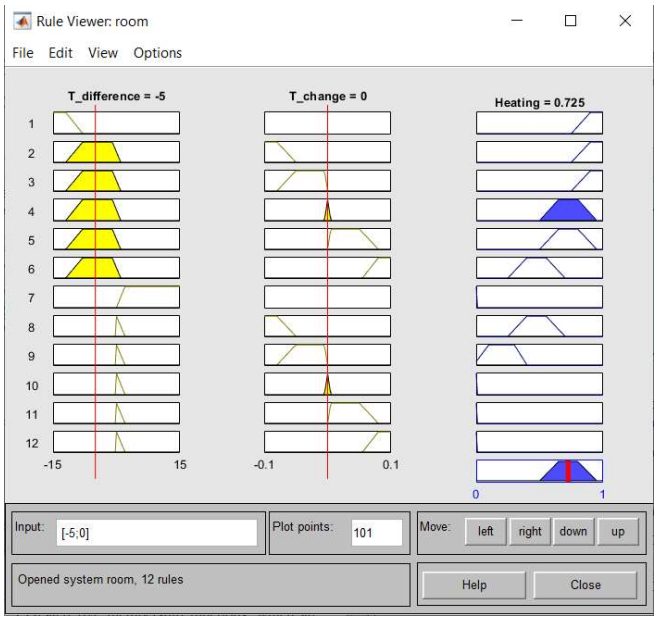


Fig. 7: Inference with fuzzy rules.

In the next step, we will demonstrate, through the Simulink model, how the prepared fuzzy controller works. Fig. 8. shows the Fuzzy control subsystem, where it can be seen, that the block has two inputs, one is the current temperature of the room, and the other is the setpoint. The "Derivative" block labeled Δu/Δt specifies how much the temperature changes in a unit of time, which can be an increase or decrease.

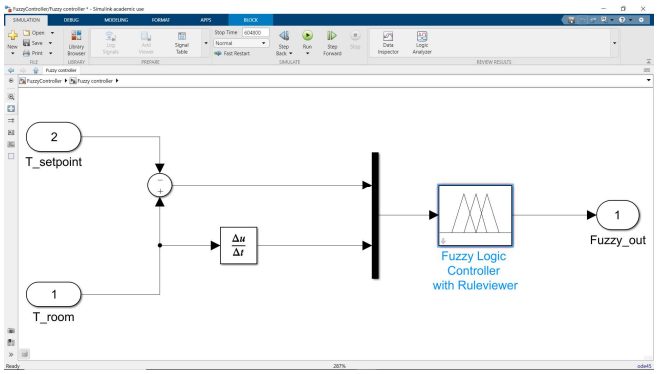


Fig. 8: Fuzzy control subsystem

The file saved from the Matlab Fuzzy Logic Toolbox must be entered and loaded into the "Fuzzy Logic Controller with Ruleviewer" block. The block has an output, which is a two-element vector containing "T_difference" and "T_change", its output value is the "Outport" block named "Fuzzy_out".

The operation of the Fuzzy controller is illustrated in our Simulink model. In the first case study, the setpoint was set at 20 °C, with a daily average of 0 °C and a daily variation of +/- 5 °C. The results of the simulation are shown in Fig. 9. (where the temperature variation of "T_room" is also zoomed in). The values displayed in Data Inspector in Fig. 10, where the setpoint, the room temperature, the outdoor temperature and the controller intervention values are shown together. From the figures, it can be seen that the room temperature maintains its setpoint with minimal fluctuation.

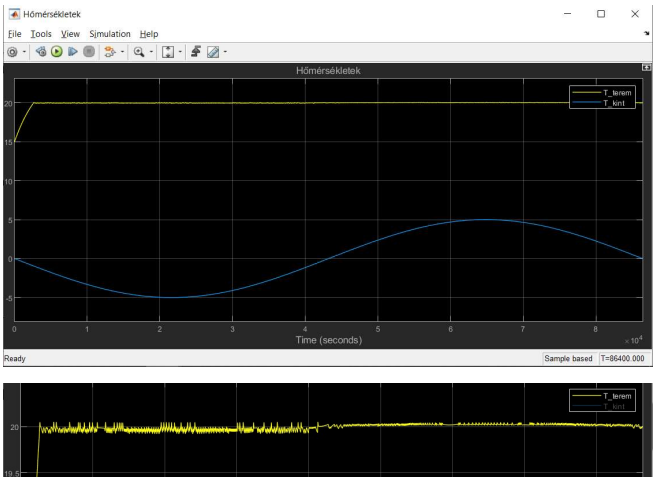


Fig. 9: Fuzzy controller operation

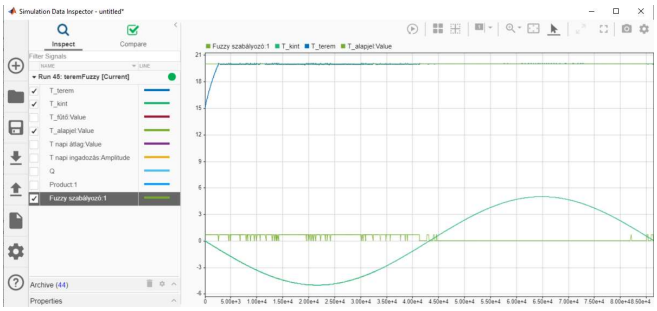


Fig. 10: Fuzzy controller operation – Data Inspector

In the second case study, we tested the setpoint change from 20 °C to 25 °C, leaving all other parameters unchanged. Here we observed how the room temperature changes when the setpoint is changed. How much time is needed for it to adjust to the new setpoint. Fig. 11. shows that the initial room temperature of 20 °C can be reached relatively quickly, but when the baseline was set to 25 °C, the system needed more time to reach the desired temperature.

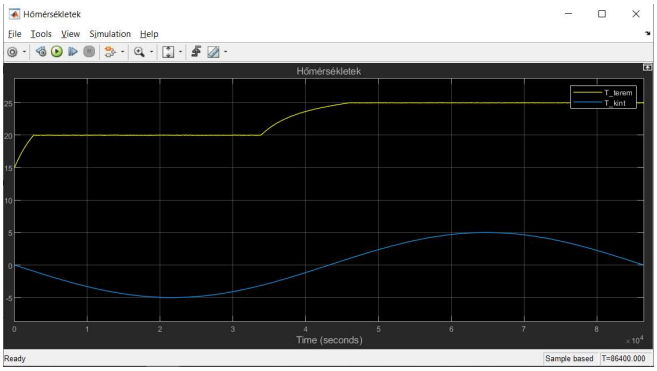


Fig. 11: Fuzzy controller operation – setpoint change

D. Comparing the fuzzy controller with other controllers

The implemented fuzzy controller was compared with two other controllers, an ON/OFF regulator and a PID controller. For both controllers, we examined the two case studies presented for the fuzzy controller.

The result of the first case study using the ON/OFF regulator can be seen in Fig. 12. and Fig. 13.

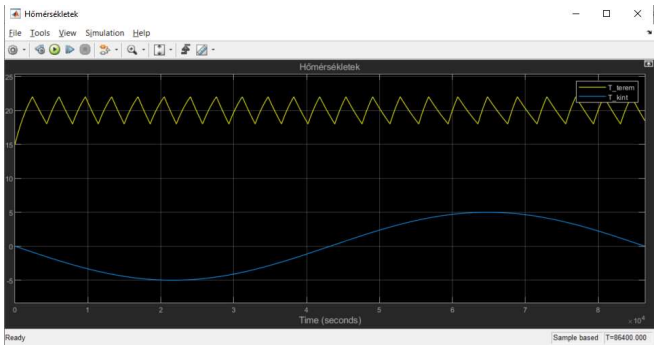


Fig. 12: Operation of the ON/OFF regulator

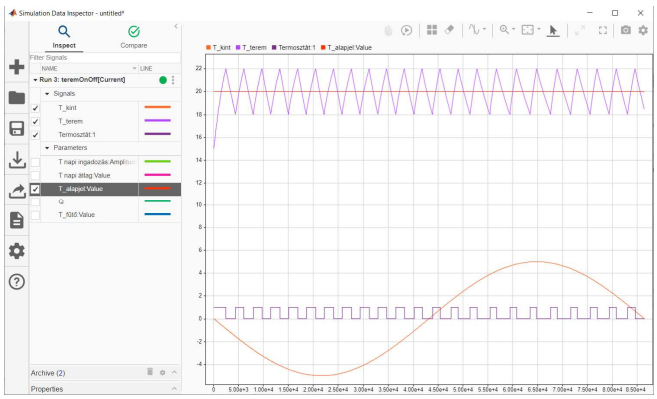


Fig. 13: Operation of the ON/OFF regulator – Data Inspector

Fig. 12 and Fig. 13. show that the setpoint of the room temperature is 20 °C, the outside temperature varies between -5 and +5 °C and the thermostat is set to switch on and off (on at 1, off at 0). It can be observed that the thermostat is on for a shorter time when the outside temperature is higher.

The second case study, i.e. the effect of the setpoint change, is shown in Fig. 14. in case of ON/OFF regulator.

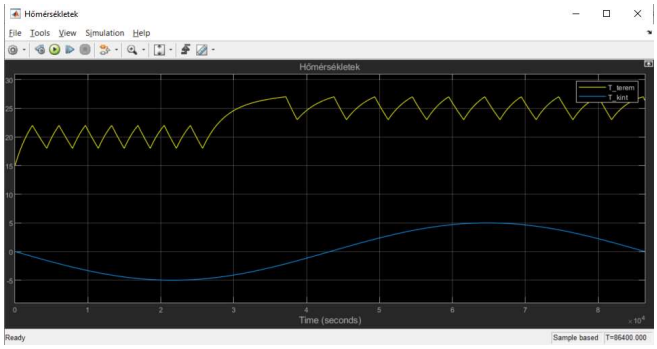


Fig. 14: Operation of the ON/OFF regulator – setpoint change

In the following, we present the heating control system with PID controller, that is the most frequently used controller in practice. Even though the actual control law is very simple, the selection of suitable parameters is not trivial. There are several methods for tuning these, but they require knowledge of control theory. The time diagram of the room and outside temperature values in the first case study is shown in Fig. 15. and Fig. 16. It can be seen that the room temperature follows the setpoint nicely after the initial fluctuation.

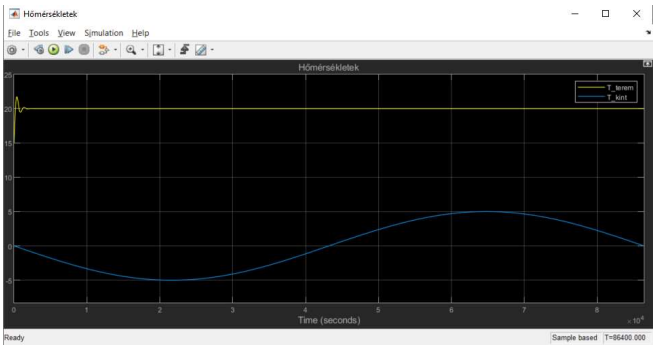


Fig. 15: PID controller operation

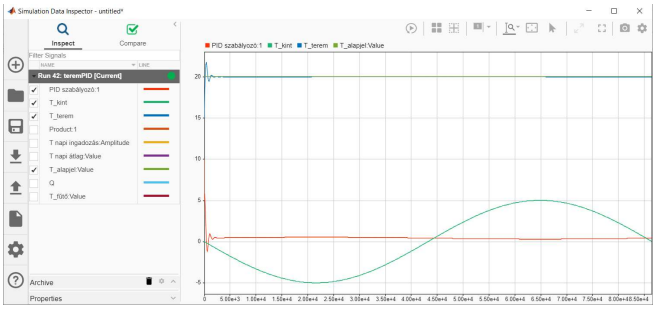


Fig. 16: PID controller operation – Data Inspector

Fig. 16. shows that the output signal of the regulator (red signal) is not limited, which is why it initially shows both "pulsed" heating and cooling. To eliminate these, the output signal value was limited to an interval [0,1] (which was the same as the ON/OFF controller limit). The result of the simulation is shown in Fig. 17.

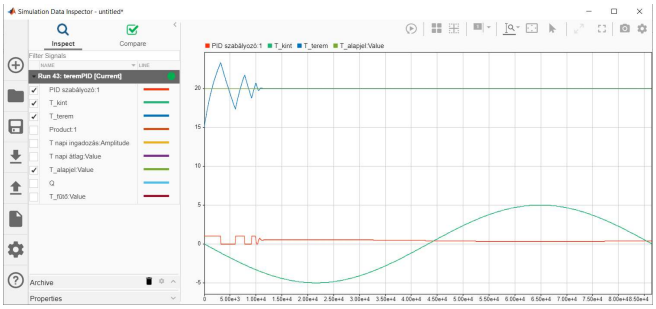


Fig. 17: PID controller operation with limits – Data Inspector

The second case study is the change of the setpoint from 20 ºC to 25 ºC, simulated in Fig. 18.

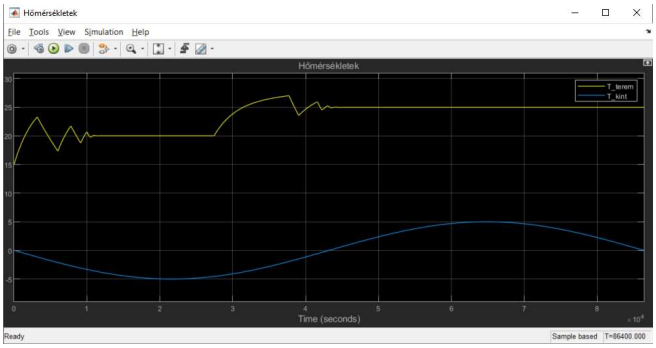


Fig. 18: PID controller operation – setpoint change

IV. CONCLUSIONS

It can be concluded that the temperature control of a room can be easily solved with all three tested regulators.

The ON/OFF regulator keeps the temperature close to the setpoint according to its operating principle. Near the setpoint, the smaller the specified error interval is, the more often it switches on. It follows the change of setpoint and the change of external temperature well. The great advantage of the regulator is its simplicity and schedulable switching function. The drawback is that it only uses two states that define full power and power off.

With the PID controller, when the setpoint is changed, the signal has a swing at first, there is an overshoot, but it always approaches the setpoint until it reaches it. With good parameter settings, it can be operated very efficiently. Several methods can be used to adjust these, but they assume knowledge of control theory.

The created Fuzzy controller was similar to the PID controller. Since we used the difference between the room temperature and the setpoint, as well as the change in the room temperature, as the controller's input, we implemented a controller with P and D members using fuzzy logic tools. The implemented controller followed both the change of the setpoint and the external temperature well. A big advantage was that the operation of the system could be described in the form of rules like human thinking, which could be implemented even with minimal control theory knowledge.

V. REFERENCES

[1] L. A. Zadeh, „Fuzzy sets," Information and Control, 8, 3, 1965.

[2] L. T. Kóczy és D. Tikk, "Fuzzy rendszerek", Budapest: TypoTEX, 2000.

[3] E. Peasley, "An Introduction to Using Simulink", Department of Engineering Science, University of Oxford , 2018.

[4] J.-S. R. Jang és N. Gulley, „The Fuzzy Logic Toolbox for use with MATLAB", Natick, Massachusetts: The MathWorks, Inc., 1995.

[5] I. S. Kobersi, V. I. Finaev, S. A. Almasani és K. W. A. Abdo, „Control of the Heating System with Fuzzy Logic," World Applied Sciences Journal, 23, 11, pp. 1441-1447, 2013.

[6] A. Taneva, D. Atanasova és A. Daskalov, „Fuzzy Logic Control in Building Automation Application," Engineering Proceedings, 41, 14, p. 10, 2023.

Smart and Increased Security Alarm System with Intrusion Detection

Bartha Csaba Richárd
Obuda University
Alba Regia Faculty
Székesfehérvár, Hungary
bartha.csaba@stud.uni-obuda.hu

Bertalan Beszédes
Obuda University
Alba Regia Faculty
Székesfehérvár, Hungary
beszedes.bertalan@uni-obuda.hu
https://orcid.org/0000-0002-9350-1802

***Abstract*—This study aims to present a smart alarm system that offers more services than the simple but expensive systems available on the market. It can be developed by users with a smaller financial investment by adding extra functions.**

***Keywords*— smart home, smart alarm system, intrusion sensor, fire alarm, smart switch automation, wifi attack, android hacking**

I. INTRODUCTION

Nowadays, the need for smart alarm systems is a modern and efficient way to protect our home and property. These systems can identify and report potential dangers, such as fire, burglary, or carbon monoxide leakage, using intelligent sensors and associated devices [1].

Smart alarm systems can be controlled remotely, allowing us to monitor the security of our home anytime, anywhere via a smartphone application [2]. These systems can send quick alerts to the appropriate authorities or the users when events are detected [3]. Intelligent technologies, such as motion detection or face and voice recognition, further enhance the efficiency of these systems [4]. Additionally, smart alarm systems can collect data on security trends and events, helping to predict and prevent potential problems [5].

Overall, these systems contribute to ensuring that our homes and their inhabitants feel safe in the modern world.

II. THEORETICAL BACKGROUND

The theoretical background of smart alarm systems is closely intertwined with the development of electronics and communication technologies. These systems use intelligent sensors and devices to identify and report potential dangers, such as fire, burglary, or carbon monoxide leakage [6].

The goal of smart alarm systems is to balance security and convenience, allowing users to monitor and control the security of their home or business remotely. These systems are often connected to the internet, enabling users to monitor their home's security live through smartphones or other devices [7].

Smart alarm systems can automatically send alerts to the appropriate authorities or users when suspicious activity is detected. Such systems typically include various sensors, such as motion sensors, door and window sensors, smoke and carbon monoxide detectors, etc. Smart alarm systems play an important role in prevention, as they can provide early warnings of potential dangers. These systems often have sophisticated, cloud-based control and monitoring platforms that allow users to access security-related information about their home from anywhere at any time [8].

Smart alarm systems are generally customizable and expandable, allowing them to adapt to unique needs and changes in the home or business environment. These systems can usually be integrated with other smart home devices, such as smart lighting or smart locks, to provide comprehensive home automation. Smart alarm systems also offer advanced data collection and analysis capabilities, helping users better understand security trends and events affecting their home [9].

Such systems typically have intuitive user interfaces that are easy and efficient to use even for non-technical users. Smart alarm systems often work independently but can cooperate with other home or business automation systems to enhance security and convenience further [10].

III. CURRENTLY AVAILABLE SOLUTIONS ON THE MARKET

There are numerous solutions on the market in the field of smart alarm systems that allow users to monitor and control their home security remotely.

A. Some popular types and services of smart alarm systems

Many companies offer additional devices that can upgrade traditional alarm systems to smart ones. These typically connect to the Wi-Fi network, allowing users to access them remotely via mobile applications or online platforms [11].

Smart home security systems include devices such as smart cameras, motion sensors, door and window sensors. These devices detect movement and send notifications to users, allowing remote monitoring [12].

Many smart alarm systems are integrated with cloud-based services. This allows storing events and data in the cloud and enables users to access and manage their data from any device [13].

Smart alarm systems are generally accessible through mobile applications. These applications allow users to arm and disarm the alarm system remotely and monitor events.

Smart alarm systems are often capable of cooperating with other smart home devices, such as smart lighting, smart locks, or smart thermostats. This allows users to achieve comprehensive home automation.

It is essential to ensure that the chosen system meets individual needs and expectations. It is always worth thoroughly reviewing the products and consulting with experts or those who have already used similar systems if necessary.

B. Purchase prices and differences in unique functionalities

Smart alarm systems are often offered in different packages. Basic packages generally include fewer sensors and basic services, while extended packages contain more sensors, cameras, and advanced functions. The type of package significantly affects the price.

The more sensors, cameras, and other smart devices are included in the system, the higher the price can be. The increase in price is usually proportional to the expanded functionality and coverage of the system.

If the system offers cloud-based services, such as storing events in the cloud or remote access, this can also affect the

price. Cloud-based services help ensure the secure storage and accessibility of your data.

If the smart alarm system is easily integrable with other smart home devices, such as smart locks, smart lighting, or smart thermostats, the price may increase. However, such integrations can result in more comprehensive home automation.

Smart alarm systems are generally accessible through mobile applications. If the applications are modern and offer comprehensive control and notifications, this can also increase the price.

C. *Interface options for smart alarm systems*

Smart alarm systems generally come with mobile applications that allow users to monitor and control the system remotely. Applications are usually available on iOS and Android platforms. Many smart alarm systems also have web interfaces, enabling management on desktop or laptop computers. Some smart alarm systems support voice control, allowing users to control the system with voice commands. Certain types of smart alarm systems may include a touchscreen panel or physical control unit for on-site management.

Systems can send email or SMS notifications to users about events detected by sensors. Smart alarm systems often connect to the cloud, allowing users to remotely monitor events and configure the system from any internet-connected device.

Smart alarm systems are generally capable of integrating with other smart home devices, such as smart lighting, smart locks, or smart thermostats. Special interfaces are available for installers and professionals for the installation, configuration, and maintenance of the system. Combining these interfaces, smart alarm systems allow users to manage the system conveniently and efficiently and monitor their home's security.

IV. SMART ALARM CENTER

A. *Structure of the smart alarm center*

SONOFF relays operate on a 230V mains power supply, connected with each other by wires as shown in the diagram below (Fig. 1.). The ultrasonic motion sensor is wired to the relay that performs the motion detection function. The relays wirelessly connect to the local network router, allowing remote control via the eWeLink mobile application. The relay performing the siren function wirelessly connects to the network and is wired to the siren.

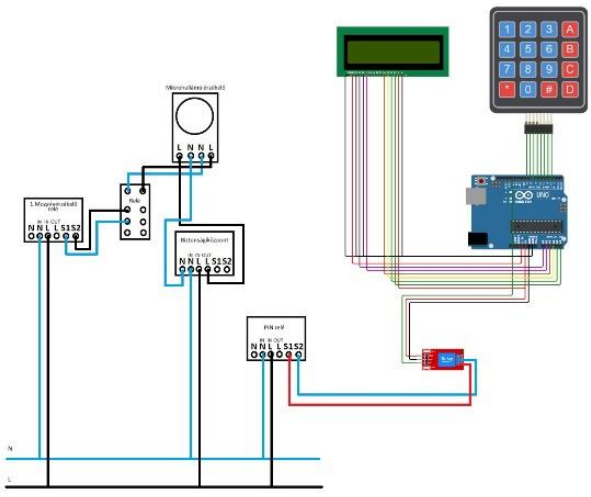


Figure 1: The schematic wiring diagram of the system

When the motion sensor detects movement, the relay performing the motion detection function is activated. When the motion detection relay is ON, and the Security System relay is also ON, the siren relay is activated, and the siren sounds. When the Security System relay is OFF, the motion detection relay loses its power supply, so the motion sensor cannot activate the motion detection relay upon detecting movement. The main flowchart can be seen on Fig. 2.

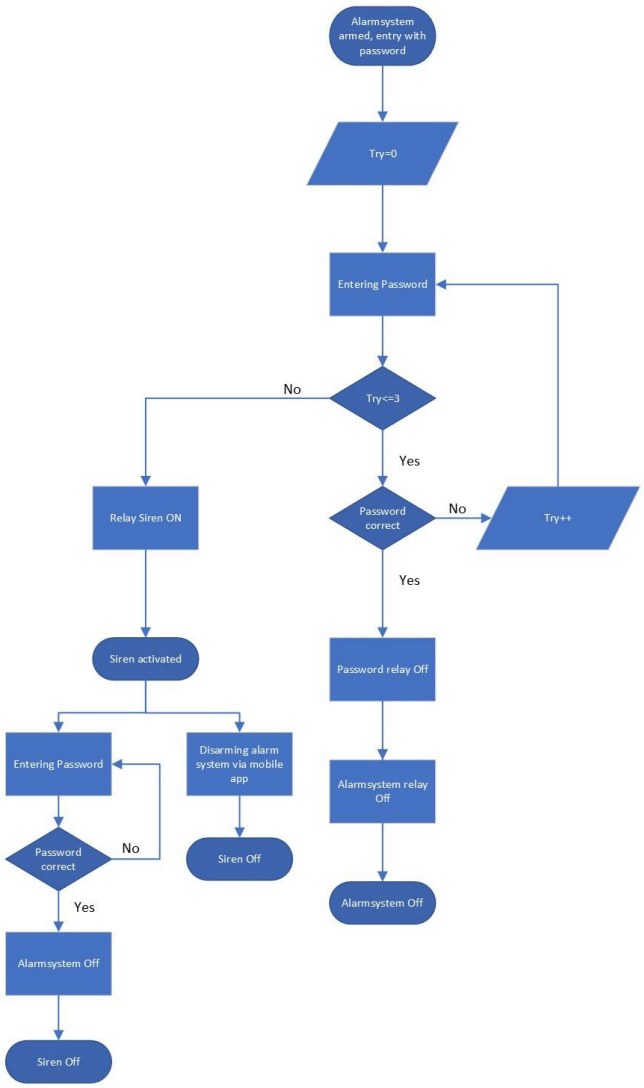


Figure 2: Flowchart of the alarm system

B. *Arming the alarm system*

Implementing a program code that allows processing, memorizing/storing a given password/code. For PIN code access, I used an Arduino Uno, an LCD display, and a membrane keypad.

The Security System (relay) is not armed (OFF):

After entering the correct PIN code, the display shows "ACCESS GRANTED press # to close", and pressing the # button activates the relay, arming the system, thereby turning on the SONOFF (PIN) wireless relay. The activated PIN relay turns on the SONOFF (security center) relay.

In the eWeLink application installed on the mobile phone, set the "Security Center" switch to ON, and the pre-configured smart control will also turn on the PIN relay (Fig. 3).

If motion is detected, the motion detection relay switches to ON, and the pre-configured smart control will also turn on the siren relay, causing the siren to sound (Fig. 4.).

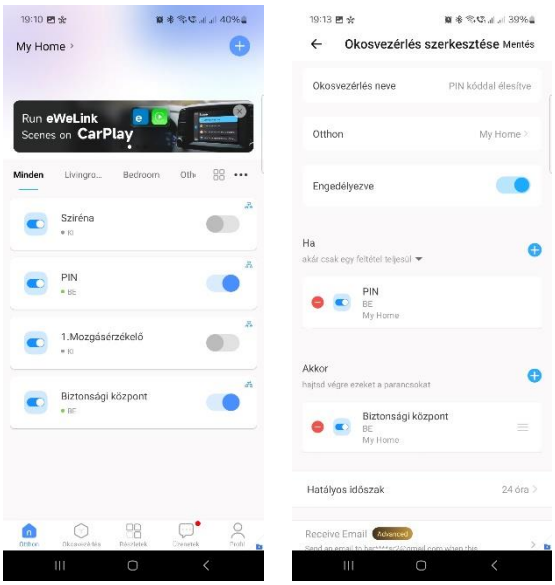


Figure 3: Setting up the application of the alarm system

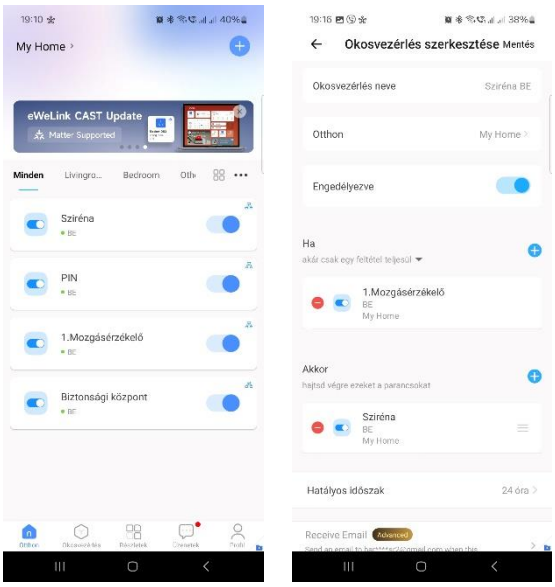


Figure 4: Modifying and testing the alarm system

C. *Disarming the alarm system can be done in two ways*

Entering the correct password disarms the alarm system (disarming with a keypad). There are 30 seconds to enter the correct code and three attempts; if time runs out or three unsuccessful attempts occur, the system locks, indicating a "burglary," the siren relay switches to ON, and the siren sounds.

Switching the Security Center to OFF in the application also disarms the system.

The application provides real-time information about all Wi-Fi relays and notifies our mobile phone of any status changes. In the mobile phone, we can link the application with the Alexa Assistant and Google Home Assistant applications.

The prototype contains the following modules: Arduino Uno, 16x2 Character LCD display module with blue backlight, 4x4 membrane keypad, Sonoff mini switch, Optonica microwave motion sensor and a 12V siren, see at Fig. 5.

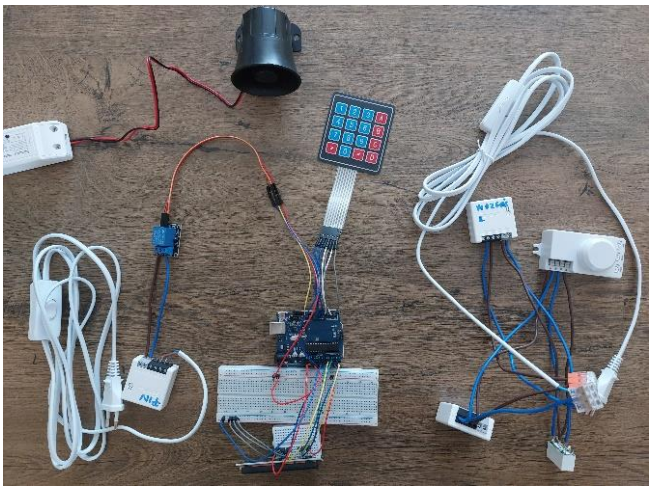


Figure 5: Prototype of the alarm system

V. THEORETICAL BACKGROUND AND SECURITY ISSUES OF SMART ALARM SYSTEMS

Smart alarm systems are based on various sensors that monitor the home environment and send alerts to the user or a security service when suspicious activity is detected. These systems typically consists sensors, central unit and communication interfaces [14-17].

Motion sensors, door and window opening sensors, smoke and heat detectors can be used. The central unit is the brain that collects data from the sensors and sends alerts. The communication module ensure data transmission between the sensors and the central unit, as well as remote access for users (e.g., via a smartphone application).

Similar hardware close, microcontroller based firmware solutions can be seen in [18-21]. The microcontroller-based solutions presented here can also be applied well in technical frontier areas [22, 23].

A. *Sabotage possibilities and risks*

The physical sabotage is a physically damaging or anunauthorised access to the alarm system devices. Software sabotage is hacking the system software or installing malicious software.

As a Protective measures, using sabotage sensors can detect physical manipulation or attempts. Secure installation also can help, it is about placing devices in hard-to-reach locations and physically protecting them. Regular software updates and applying security patches and implementing multiple security mechanisms, such as firewalls, antivirus programs, and intrusion detection systems will increase the resistence against software sabotage.

Communication between all elements of the smart alarm system must be reliable and secure. It is proposed to regularly checking the system to ensure all elements function correctly, also setting up automatic notifications and alerts for errors, intrusion attempts, or other suspicious activities [24, 25]. The proposed solution is well suitable for other applications in the term of safety data acquisition [26, 27].

B. *Loss of connection between modules*

The loss of connection between modules in the alarm system presents several risks. Firstly, the efficiency of the system may decrease if the connection between different modules is lost. This can lead to significant notification gaps,

where alerts may not reach the user or the security service in a timely manner.

To mitigate these risks, several protective measures can be implemented. Utilizing redundant communication channels, such as Wi-Fi and mobile networks, ensures that if one channel fails, the other can maintain the connection. Additionally, conducting regular automatic system checks can immediately notify the user of any connection loss. Setting up outage alerts can also provide warnings about connection loss or communication problems between modules, helping to maintain the overall reliability and effectiveness of the alarm system.

C. Smart devices associated with the system's security

The security of smart devices like phones, laptops, Wi-Fi networks, and assistant services is crucial, as intruders can potentially access the alarm system through these devices. Several risks are associated with compromised smart devices. Firstly, there is the risk of personal data compromise, where intruders can steal or manipulate personal data stored on these devices. Secondly, there is the risk of remote control compromise, where hackers can take control of smart devices remotely.

To protect against these risks, several measures should be implemented. Ensuring device security involves setting up safe configurations for smart devices, such as using screen locks, strong passwords, and regularly updating software. Protecting the Wi-Fi network with strong passwords and encryption is also essential. Additionally, smart assistant services should be properly configured with appropriate access permissions and privacy settings to enhance security.

D. Major security risks

Several major security risks threaten smart home security systems. One significant risk is signal jamming, where hackers can jam wireless signals, interrupting communication between sensors and the control panel. This can significantly compromise the effectiveness of the security system.

Another critical risk is network access. If a hacker gains access to the Wi-Fi network, they can control connected smart devices, including cameras and smart locks, posing a severe security threat.

Device hacking is also a major concern. Poorly protected devices, especially those using default or weak passwords, can be easily hacked. This allows intruders to control the devices remotely and bypass security measures.

Lastly, software and firmware exploitation poses a significant risk. Many smart devices run outdated software or firmware with unpatched security vulnerabilities. Hackers can exploit these weaknesses to gain control over the devices, further compromising the security system.

E. Protective measures

Each device should have strong, unique passwords, means avoid default passwords and change them regularly. It is possible to use a password manager to handle complex passwords.

Using WPA2 or WPA3 encryption on the router secures the Wi-Fi network. Recommended to set a strong Wi-Fi password and considering creating a separate network for smart devices to isolate them from other devices.

Proposed to keep all smart devices updated with the latest software and firmware updates to patch security vulnerabilities and enable automatic updates if possible.

Enable two-factor authentication for smart home applications and accounts adds an extra layer of security in addition to the password.

It is proposed to regularly check the network for unauthorized devices and to use network management tools to see what devices are connected and remove unknown ones.

Good practice to purchase smart devices from reputable manufacturers who prioritize security and provide regular updates. Also to research and read reviews to ensure the devices have a good security track record.

VI. PRACTICAL IMPLEMENTATION

In the project, the Sonoff relays and the microwave motion sensor power wires are wired to each other. The router provides access to the network through a wireless connection with WPA2 encryption. In case of internet outage, the system cannot be controlled remotely; it continues to operate within the local network, and the application on the remote device will indicate offline status. The armed system can still be disarmed with the Arduino Uno-controlled keypad using a PIN code, and if the mobile device connects to the router's wireless network, it can also be disarmed from the application.

In the event of a router "freeze," using the keypad ensures the on-site activation and deactivation of the system.

The access and privacy settings of the assistant on the mobile device have been configured to avoid hacker attacks.

To calculate the annual energy consumption, check the consumption of the following devices:

1) Sonoff Mini relay:
   - Standby consumption: 0.5W
   - Operating consumption: 1.5W (maximum)
2) Optonica microwave motion sensor:
   - Consumption: approx. 0.5W
3) Arduino Uno:
   - Average consumption: approx. 0.5W

A. Annual consumption

Consumption for all devices in standby mode:

- Sonoff Mini (3 pcs): 1.5W
- Optonica motion sensor: 0.5W
- Arduino Uno: 0.5W

The total consumption in standby mode is:

$$1.5\text{W} + 0.5\text{W} + 0.5\text{W} = 2.5\text{W} \quad (1)$$

Consumption for all devices in operating mode:

- Sonoff Mini (3 pcs): 4.5W
- Optonica motion sensor: 0.5W
- Arduino Uno: 0.5W

The total consumption in operating mode is:

$$4.5\text{W} + 0.5\text{W} + 0.5\text{W} = 5.5\text{W} \quad (2)$$

Annual energy consumption in standby mode, based on formula (1):

$$2.5\text{W x } 24 \text{ hours x } 365 \text{ days} = 21.9\text{kWh} \quad (3)$$

Annual energy consumption in standby mode, based on formula (2):

$$5.5\text{W x } 24 \text{ hours x } 365 \text{ days} = 48.18\text{kWh} \quad (4)$$

B. Annual cost

If the local electricity price is approx. 38 HUF/kWh, then the annual cost in standby mode, based on formula (3):

$$21.9\text{kWh x } 38 \text{ HUF/kWh} = 832 \text{ HUF} \quad (5)$$

and in operating mode, based on formula (4):

$$48.18\text{kWh x } 38 \text{ HUF/kWh} = 1830 \text{ HUF} \quad (6)$$

In an average household, the estimated average annual consumption is: 45kWh = 1710 HUF

This summary shows that the annual energy consumption and cost of three Sonoff Mini relays, one Optonica microwave motion sensor, and one Arduino Uno are low, providing an energy-efficient solution for smart homes.

C. System uptime in case of power outage using an uninterruptible power supply

To calculate uptime during a power outage, consider the system's energy consumption and the UPS battery capacity.The system uses a CyberPower BU650E uninterruptible power supply with a capacity of 12V 7Ah, which equals approximately 84 Wh.

In standby mode, the total consumption is 2.5 W, resulting in an uptime of 33.6 hours (calculated as 84Wh divided by 2.5W). In operating mode, the total consumption is 5.5 W, resulting in an uptime of 15.27 hours (calculated as 84Wh divided by 5.5W).

For summary, the CyberPower BU650E can typically operate the system for approximately 30-35 hours during a power outage.

CONCLUSION

In conclusion, this study successfully demonstrates the potential of a smart alarm system that surpasses the functionalities of existing market offerings while remaining financially accessible. By enabling users to customize and expand the system with additional features, this innovative approach addresses both cost-efficiency and functionality.

The proposed system not only serves as a viable alternative to expensive commercial options but also empowers users to tailor their security solutions according to their specific needs.

Future work will focus on refining the system's capabilities, enhancing user-friendliness, and conducting extensive field testing to ensure reliability and effectiveness. This development marks a significant step forward in making advanced security systems more accessible and versatile for a broader audience.

ACKNOWLEDGMENT

The author would like to thank all the Obuda University Alba Regia Faculty staff and member that provide help and assistance throughout the project completion.

REFERENCES

[1] Javale, Deepali, et al. Home automation and security system using Android ADK. International journal of electronics communication and computer technology (IJECCT) 3.2 (2013): 382-385.

[2] Chen, Joy Iong Zong. "Smart security system for suspicious activity detection in volatile areas." Journal of Information Technology 2.01 (2020): 64-72.

[3] VIDYAVIHAR, MUMBAI. "Remote Controlled Home Automation Using Android Application via WiFi Connectivity." International Journal on Recent and Innovation Trends in Computing and Communication 3.3 (2015): 1489-1492.

[4] Shah, Rakshit, et al. "Wireless smoke detector and fire alarm system." International Research Journal of Engineering and Technology (IRJET) 6.1 (2019): 1407-1412.

[5] Bertalan, Beszédes. Reliable Presence and Intrusion Detection with Collaborative Sensor Modules in Electronic Property Protection Systems. In: IEEE (szerk.) 2019 IEEE 17TH World Symposium on Applied Machine Intelligence and Informatics (SAMI 2019). Herlany, Szlovákia : IEEE (2019) pp. 354-360. , 7 p.

[6] Abdelouahid, R.A., Debauche, O., Marzak, A.: Internet of things: a new interoperable IoT platform. Application to a smart building. Proced Comput Sci 191, pp. 511–517 (2021)

[7] Vadakkan, Annmary, et al. "Door locking using keypad and ARDUINO." International Research Journal of Modernization in Engineering Technology and Science 3.11 (2021): 780-787.

[8] Orji, E. Z., U. I. Nduanya, and C. V. Oleka. "Microcontroller Based Digital Door Lock Security System Using Keypad." International Journal of Latest Technology in Engineering, Management & Applied Science 8.1 (2019): 92-97.

[9] George, Abin. "Automation Of Main Gate, Garage Door and Washing System."

[10] Iqbal, Muhammad Javed, et al. "Smart home automation using intelligent electricity dispatch." Ieee Access 9 (2021): 118077-118086.

[11] Aung, May Aye Chan, and Khin Phyo Thant. IEEE 802.11 attacks and defenses. Diss. MERAL Portal, 2019.

[12] Daniyar Kaliyev, Olga Shvets, György Györök. Computer Vision-based Fire Detection using Enhanced Chromatic Segmentation and Optical Flow Model. Acta Polytechnica Hungarica Vol. 20, No. 6, 2023. ISSN: 1785-8860

[13] Misra, Anmol, and Abhishek Dubey. Android security: attacks and defenses. CRC Press, 2013.

[14] Mosavi Amirhosein, Bertalan Beszedes, Imre Felde, Nadai Laszlo, Gorji Nima E. Electrical characterization of CIGS thin-film solar cells by two- and four-wire probe technique. MODERN PHYSICS LETTERS B 2020 p. 2050102 , 16 p. (2020)

[15] György Györök, Baklanov Alexander E, Bertalan Beszédes. Extension of Nodal Voltage Method with the Thermosensing. In: Orosz, Gábor Tamás (szerk.) AIS 2017 - 12th International Symposium on Applied Informatics and Related Areas organized in the frame of Hungarian Science Festival 2017 by Óbuda University : Proceedings. Székesfehérvár, Magyarország : Óbudai Egyetem (2017) 204 p. pp. 201-204. , 4 p.

[16] György Györök, András Dávid, Nikoletta Tolner, Bertalan Beszédes, Dániel Cseh. Supervision of the operation of digital circuits by Embedded Microcontroller. In: Orosz, Gábor (szerk.) AIS 2018 - 13th International Symposium on Applied Informatics and Related Areas. Székesfehérvár, Magyarország : Óbudai Egyetem, Alba Regia Műszaki Kar (2018) pp. 29-35. , 7 p.

[17] Györök György, Tihomir Trifonov, Alexander E. Baklanov, Bertalan Beszédes, Svetlana V. Grigoryeva, Aizhan Zhaparova. A Special Robust Solution for Battery Based Power Supply. In: Orosz, Gábor Tamás (szerk.) 11th International Symposium on Applied Informatics and Related Areas (AIS 2016). Székesfehérvár, Magyarország : Óbudai Egyetem (2016) pp. 32-35. , 4 p.

[18] Attila Sáfár, Bertalan Beszédes. Educational Aspects of a Modular Power Management System. In: Orosz, Gábor Tamás (szerk.) AIS 2019 : 14th International Symposium on Applied Informatics and Related Areas organized in the frame of Hungarian Science Festival 2019 by Óbuda University. Székesfehérvár, Magyarország : Óbudai Egyetem (2019) pp. 163-166. , 4 p.

[19] György Györök, Bertalan Beszédes. Artificial Education Process Environment for Embedded Systems. In: Orosz, Gábor Tamás (szerk.) 9th International Symposium on Applied Informatics and Related

Areas - AIS2014. Székesfehérvár, Magyarország : Óbudai Egyetem (2014) pp. 37-42. , 6 p.

[20] Alexander Baklanov, Svetlana Grigoryeva, György Györök. Control of LED Lighting Equipment with Robustness Elements, Acta Polytechnica Hungarica (1785-8860 1785-8860): 15 3 pp 105-119 (2016)

[21] G. Gyorok and M. Mako, "Configuration of EEG input-unit by electric circuit evolution," 2005 IEEE International Conference on Intelligent Engineering Systems, 2005. INES '05., Spain, 2005, pp. 255-258, doi: 10.1109/INES.2005.1555168.

[22] Aizhan Zhaparova, Dimitry Titov, Alexander Y. Baklanov, Gyorgy Gyorok. Study of the Effectiveness of Switching-on LED Illumination Devices and the Use of Low Voltage System in Lighting. Acta Polytechnica Hungarica (1785-8860 1785-8860): 12 5 pp 71-80 (2015)

[23] Thermo-dynamic Cycle Computation of a Micro Turbojet Engine. Fozo, L; Andoga, R and Kovacs, R. 17th IEEE International Symposium on Computational Intelligence and Informatics (CINTI). 2016 17TH IEEE International Symposium on Computational Intelligence and Informatics (CINTI 2016) , pp.75-79

[24] Ediboglu Bartos Gaye, Akyol Serel. Deep Learning for Image Authentication: A Comparative Study on Real and AI-Generated Image Classification. In: Petőné Csuka, Ildikó; Simon, Gyula (szerk.) AIS 2023 - 18th International Symposium on Applied Informatics and Related Areas – Proceedings. Székesfehérvár, Magyarország : Óbudai Egyetem (2023) 177 p. pp. 23-27. 5 p.

[25] Majnár Csaba, Rádli Ferenc Gábor, Ediboglu Bartos Gaye. Augmented Reality in High School Literature: Enhancing Engagement and Evaluation. In: Petőné Csuka, Ildikó; Simon, Gyula (szerk.) AIS 2023 - 18th International Symposium on Applied Informatics and Related Areas – Proceedings. Székesfehérvár, Magyarország : Óbudai Egyetem (2023) 177 p. pp. 155-158. 4 p.

[26] T., Jancso ; A., Podor ; E., Nagyne Hajnal ; P., Udvardy ; G., Nagy ; A., Varga ; Q., Meng ; L., Zhang. Data Integration with Geographic Information System Tools for Rural Environmental Monitoring. In: IRC 2022 XVI. International Research Conference Proceedings. Peking, Kína (2022) 270 p. pp. 52-59. , 8 p.

[27] Peter, Udvardy; Levente, Dimen; Gergely, Vakulya. Data Collection In The Wild: Challenges And Solutions. University of Agronomic Sciences and Veterinary Medicine of Bucharest. Scientific Papers. Series E. Land Reclamation, Earth Observation And Surveying, Environmental Engineering 12 pp. 167-175. , 9 p. (2023)

Embedded Control System for Residential Heating, Cooling, and Ventilation

Szabolcs Schekk
Obuda University
Alba Regia Faculty
Székesfehéhérvár, Hungary
sschekk@stud.uni-obuda.hu

Bertalan Beszédes
Obuda University
Alba Regia Faculty
Székesfehérvár, Hungary
beszedes.bertalan@uni-obuda.hu
https://orcid.org/0000-0002-9350-1802

***Abstract*—This paper discusses the development of a universal controller designed for households equipped with modern HVAC technology. The controller's task is to optimize heating, cooling, and ventilation, as well as manage the Smart Grid functionality of heat pump systems. The paper presents the hardware and software solutions employed, and provides a detailed description of the system's operation and structure.**

***Keywords*— heating control, cooling control, ventilation control, smart grid, smart control system, residential heating, residential cooling, residential ventilation**

I. INTRODUCTION

In modern households, comfortable and energy-efficient climate control is essential. In this paper, we present the development of a universal controller designed to meet these needs. The controller aims to optimize heating, cooling, and ventilation, taking into account external temperatures, user preferences, and the current state of the energy provider's infrastructure via the SmartGrid [1, 2] system. Energy efficiency is achieved through appropriate weather compensating regulation, the advantage of which lies in reducing the temperature fluctuations and inertia losses associated with traditional thermostat systems. This can result in up to 10% energy and cost savings [3, 4].

In addition to energy efficiency benefits, ventilation regulation also has favorable physiological effects. Humans perceive indoor air quality primarily through smell and via the conjunctiva of the eyes. Poor air quality can lead to poisoning in severe cases, while prolonged exposure to lower quality air can cause irritation and promote the development of chronic respiratory diseases [5-7].

II. MATERIALS AND METHODS

Residential heat pumps generally come with an internal changeover valve, which determines whether to produce hot water or heating water based on current needs, follow on Fig. 1. and check Fig. 2. The internal changeover valve directs the water flow accordingly towards hot water production or heating. The supplementary heating unit may not have an internal changeover valve, so external changeover valves are placed in the flow and return pipes. These operate in parallel with each other. Heating appliances have their own pumps. Check valves ensure that water circulates securely between heating circuits or storage tanks, without flowing through each other. The buffer tank functions as an energy storage unit [8].

When the controller determines that a sufficient amount of energy is available, it activates the buffer tank into the heating system by switching the changeover valve before the buffer tank. The hydraulic balancer ensures the necessary flow rate for the heat generator and heating circuit pumps without interfering with each other. A mixing valve control system consists of a pump, a mixing valve, and a forward water temperature sensor. Both the switching and mixing valves are three-way valves, but the difference is that the mixing valve can also be in an intermediate position, while the switching one only switches between the two end positions.

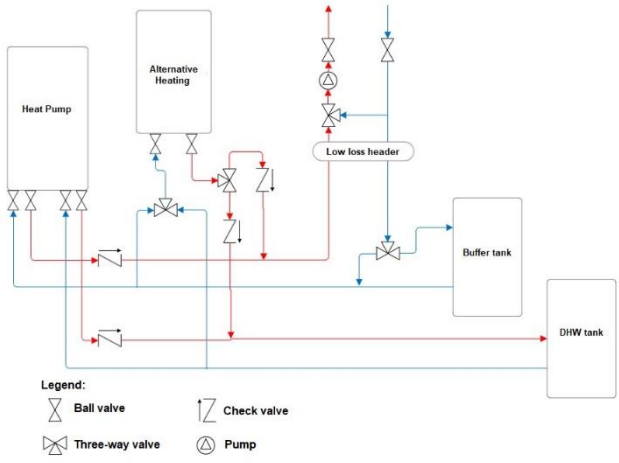


Figure 1: General hydraulic diagram of a residential heating system

Figure 2: A photograph of a heating center based of the diagram

For hardware design, selecting a suitable microcontroller with sufficient I/O ports, ADCs, and communication interfaces is essential. The Arduino UNO R4 WiFi [9, 10] card forms the basis of the control, with its built-in ESP32-S3 module enabling wireless communication. Standard heating

appliances, pumps, and control valves can be used in the system. Heating appliances must have a standard thermostat input. This is a two-pole connector. To operate this thermostat input, a potential-free contact must be connected (see on Fig. 3. and Fig. 4.). The heating appliance pulls one pole of the input to a level it can interpret (usually 230 VAC), then expects the same voltage level back on its other pole. The closed position of the contact indicates the need for heating.

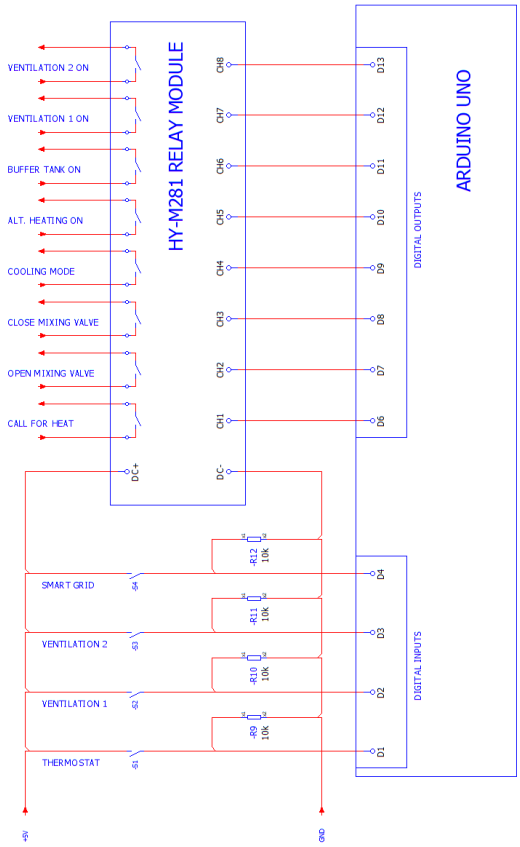


Figure 3.: Allocation of digital inputs and outputs of the controller

All standard household pumps are compatible with the device. A single-phase pump can be controlled through the controller up to 2 Amps, but it is recommended to isolate the pumps with relays or contactors as a solution. In this case, we certainly do not overload the output of the controller, and even a three-phase pump can be connected, or in the case of a single-phase pump, we can also disconnect the neutral wire.

This is a common practice in boiler rooms or particularly humid, moist rooms. Any type of changeover valve can be applied as well. In the case of a spring-return design, direct control from the controller is possible, otherwise, relay isolation is required for switching between the opening and closing directions. Only a stepper-type design can be used for the mixing valve since the controller controls the valve with a digital signal. For analog control mixing valves (0-10 V, 4-20 mA), additional accessories are required.

The temperature sensors should be of the Pt1000 type. Instead of resistance thermometers, we could have used thermocouples or standard temperature transmitters. We chose Pt1000 for its reliability, favorable price, and widespread availability [11, 12].

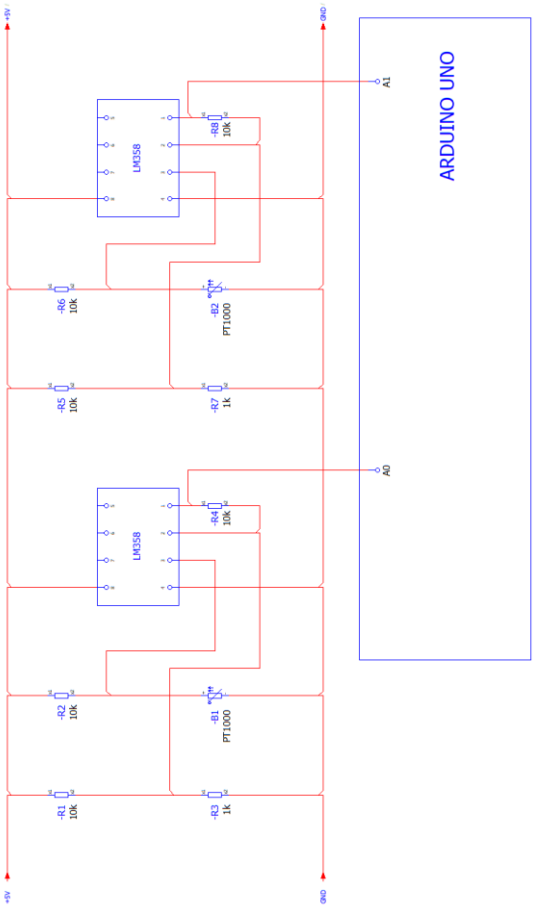


Figure 4: Temperature sensing by the controller

III. Results

The developed controller successfully manages the functions of heating, cooling, and ventilation, taking into account external environmental conditions, user preferences, and energy provider system status indicators (see on Fig. 5. and Fig. 6.). The system ensures stable operation and ensures user-friendly interface.

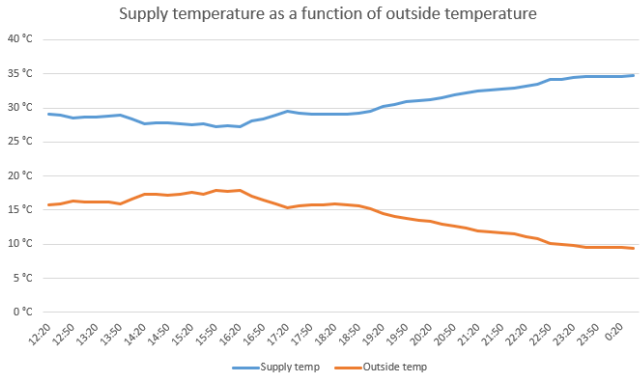


Figure 5: Graph of the measured temperature values during the operation of the controller

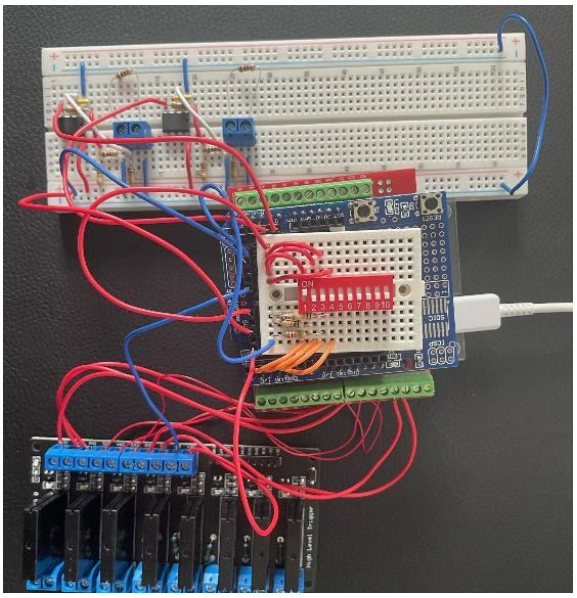


Figure 6: Controller subsystems and operation

Creating an embedded control system for residential heating, cooling, and ventilation involves designing both the hardware and software components to ensure efficient and reliable operation [13, 14]. With different sensors and actuators, it is a multiple input and multiple output system, therefore various control methods can be used [15-17]. In subprocesses control algorithms such as PID control is used for keeping up the desired technological levels. Sensor data processing subroutines includes filtering techniques to remove noise and calibration routines for accuracy [18-20].

The comprehensive software state diagram can be seen on Fig. 7. and the firmware implementation can be seen on Fig. 8. Besides that, the system contains communication interfaces and handles the user interface. Additionally, energy optimization algorithms [21-23] are incorporating to enable energy-efficient operation, adjusting scheduling and setpoints based on occupancy and external weather conditions [24-26].

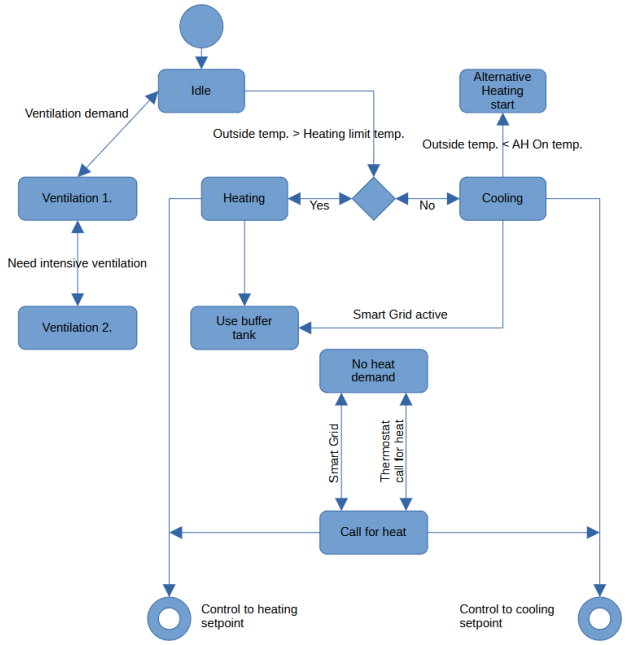


Figure 7: Controller state diagram

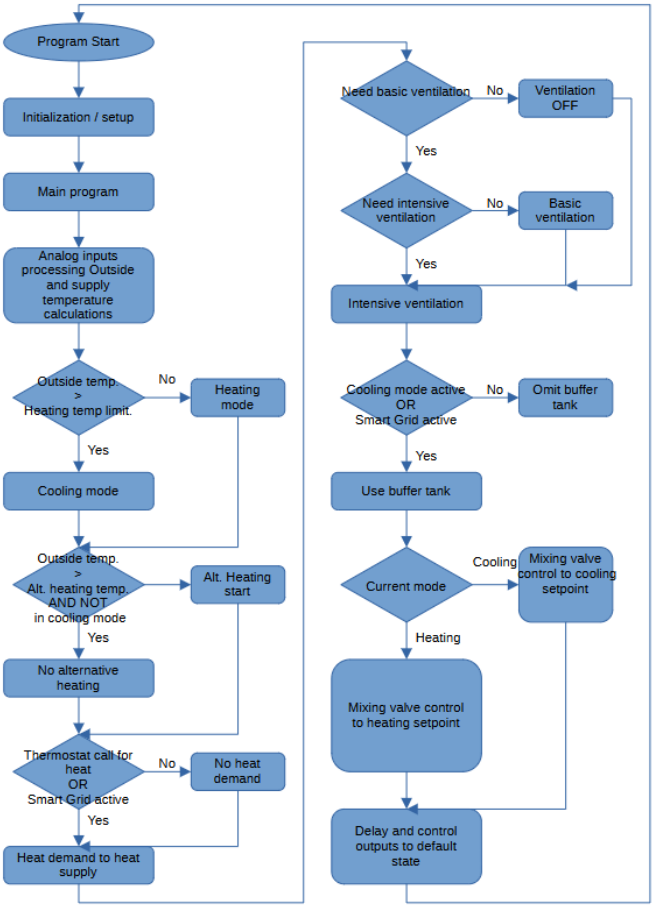


Figure 8: Controller program flowchart

Further optimization of code for efficiency, responsiveness and fine-tuning control algorithms for optimal performance are necessary in each different application before final deployment. Fors maintenance and updates for firmware and software can be done via the open source development environment, also monitoring and diagnostics, and implementing alert systems for faults and maintenance needs will help maintain system reliability and efficiency.

In the context of an embedded control system for residential climatization, effective data collection and secure data management are crucial for optimizing performance and ensuring user safety. The integration of data collection methodologies, as discussed in [27] highlights the significance of using geographic information systems for gathering environmental data in residential settings. The framework facilitates real-time data integration, allowing for a comprehensive understanding of the spatial and temporal dynamics of residential energy use, ultimately contributing to improved system efficiency. [28] also emphasize the importance of secure data storage and transmission within such embedded systems. As residential environments increasingly rely on interconnected devices, ensuring the security of sensitive data becomes paramount. Implementing robust encryption protocols and secure communication channels will safeguard user information and prevent unauthorized access. The combined insights from both publications underscore the necessity of a holistic approach to data management, where efficient data collection techniques are complemented by stringent security measures.

IV. Conclusion

The presented embedded control system for residential heating, cooling, and ventilation involves integrating various sensors and actuators with a microcontroller, developing control algorithms, and ensuring a user-friendly interface. It has resulted in an efficient and flexible system capable of meeting the climate control needs of modern households. The system in focusing on maintaining comfort, ensuring air quality, and optimizing energy consumption. The technologies and solutions employed in the development allow for future enhancements and optimizations to be implemented.

Acknowledgement

The author would like to thank all the Obuda University Alba Regia Faculty staff and member that provide help and assistance throughout the project completion.

References

[1] Fischer, D.; Wolf, T.; Triebel, M.A. Flexibility of heat pump pools: The use of SG-Ready from an aggregator's perspective. InProceedings of the 12th IEA Heat Pump Conference 2017, Rotterdam, The Netherlands, 15–18 May 2017.

[2] Rapucha, Adrian, Ramadas Narayanan, and Meena Jha. 2024. "Heat Pumps with Smart Control in Managing Australian Residential Electrical Load during Transition to Net Zero Emissions" *Energies* 17, no. 12: 2977. https://doi.org/10.3390/en17122977

[3] André Jacob. „Regulations for the "SG ready" label for electrically driven heating and hot water heat pumps and interface-compatible system components" Bundesverband Wärmepumpe (BWP). Version 2.0

[4] Márton L. Kiss, Klára Tóthné Szita, Judit Mária Pintér, Balázs Kósa, Laura Veres. Lifecycle Analysis of an Air Quality Sensor. Acta Polytechnica Hungarica. Vol. 21, No. 8, 2024. ISSN: 1785-8860

[5] World Health Organization. Environmental Health in Rural and Urban Development and Housing Unit. (1990). Indoor environment : health aspects of air quality, thermal environment, light and noise. World Health Organization. https://iris.who.int/handle/10665/62723

[6] World Health Organization. Environmental Health in Rural and Urban Development and Housing Unit. (1990). Indoor environment : health aspects of air quality, thermal environment, light and noise. World Health Organization. https://iris.who.int/handle/10665/62723

[7] Cincinelli A, Martellini T. Indoor Air Quality and Health. Int J Environ Res Public Health. 2017 Oct 25;14(11):1286. doi: 10.3390/ijerph14111286. PMID: 29068361

[8] Bjarne Olesen, Control of floor heating and cooling systems. Clima 2000/Napoli 2001 World Congress – Napoli (I), 15-18 September 2001

[9] Csilla Patkó, István Patkó, Zoltán Pásztory. Indoor Air Quality Testing in Low-Energy Wooden Houses: Measurement of Formaldehyde and VOC-s. Acta Polytechnica Hungarica. Vol. 10, No. 8, 2013. ISSN: 1785-8860

[10] Bonnie Baker. Temperature sensing technologies. Microchip Technology Inc. AN679. 2002.

[11] Mosavi Amirhosein, Bertalan Beszedes, Imre Felde, Nadai Laszlo, Gorji Nima E. Electrical characterization of CIGS thin-film solar cells by two- and four-wire probe technique. MODERN PHYSICS LETTERS B 2020 p. 2050102 , 16 p. (2020)

[12] György Györök, Baklanov Alexander E, Bertalan Beszédes. Extension of Nodal Voltage Method with the Thermosensing. In: Orosz, Gábor Tamás (szerk.) AIS 2017 - 12th International Symposium on Applied Informatics and Related Areas organized in the frame of Hungarian Science Festival 2017 by Óbuda University : Proceedings. Székesfehérvár, Magyarország : Óbudai Egyetem (2017) 204 p. pp. 201-204. , 4 p.

[13] György Györök, András Dávid, Nikoletta Tolner, Bertalan Beszédes, Dániel Cseh. Supervision of the operation of digital circuits by Embedded Microcontroller. In: Orosz, Gábor (szerk.) AIS 2018 - 13th International Symposium on Applied Informatics and Related Areas. Székesfehérvár, Magyarország : Óbudai Egyetem, Alba Regia Műszaki Kar (2018) pp. 29-35. , 7 p.

[14] Györök György, Tihomir Trifonov, Alexander E. Baklanov, Bertalan Beszédes, Svetlana V. Grigoryeva, Aizhan Zhaparova. A Special Robust Solution for Battery Based Power Supply. In: Orosz, Gábor Tamás (szerk.) 11th International Symposium on Applied Informatics and Related Areas (AIS 2016). Székesfehérvár, Magyarország : Óbudai Egyetem (2016) pp. 32-35. , 4 p.

[15] György Györök, Bertalan Beszédes. Artificial Education Process Environment for Embedded Systems. In: Orosz, Gábor Tamás (szerk.) 9th International Symposium on Applied Informatics and Related Areas - AIS2014. Székesfehérvár, Magyarország : Óbudai Egyetem (2014) pp. 37-42. , 6 p.

[16] I. Bendiák and T. Sándor, "Possible Ways of Measuring and Calculating Waste Heat from a Machine Diagnostic Approach," *2023 IEEE 6th International Conference and Workshop Óbuda on Electrical and Power Engineering (CANDO*-EPE), Budapest, Hungary, 2023, pp. 000259-000266, doi: 10.1109/CANDO-EPE60507.2023.10417972.

[17] I. Bendiák and S. Semperger, "Simplified Predictive Strategy of Mechanical Life Cycle Model in Three-Phase Asynchronous Motor," 2023 IEEE 6th International Conference and *Workshop Óbuda on Electrical and Power Engineering (CANDO-EPE)*, Budapest, Hungary, 2023, pp. 000267-000274, doi: 10.1109/CANDO-EPE60507.2023.10417977.

[18] Alexander Baklanov, Svetlana Grigoryeva, György Györök. Control of LED Lighting Equipment with Robustness Elements, Acta Polytechnica Hungarica (1785-8860 1785-8860): 15 3 pp 105-119 (2016)

[19] G. Gyorok and M. Mako, "Configuration of EEG input-unit by electric circuit evolution," 2005 IEEE International Conference on Intelligent Engineering Systems, 2005. INES '05., Spain, 2005, pp. 255-258, doi: 10.1109/INES.2005.1555168.

[20] Aizhan Zhaparova, Dimitry Titov, Alexander Y. Baklanov, Gyorgy Gyorok. Study of the Effectiveness of Switching-on LED Illumination Devices and the Use of Low Voltage System in Lighting. Acta Polytechnica Hungarica (1785-8860 1785-8860): 12 5 pp 71-80 (2015)

[21] G. Vakulya and G. Simon, "Energy-efficient and reliable round-robin TDMA for wireless sensor networks," 2012 IEEE International Instrumentation and Measurement Technology Conference Proceedings, Graz, Austria, 2012, pp. 1179-1183, doi: 10.1109/I2MTC.2012.6229240.

[22] G. Zachár, G. Vakulya and G. Simon, "Design of a VLC-based beaconing infrastructure for indoor localization applications," 2017 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), Turin, Italy, 2017, pp. 1-6, doi: 10.1109/I2MTC.2017.7969837.

[23] Simon, Gyula, Gergely Vakulya, and Márk Rátosi. 2022. "The Way to Modern Shutter Speed Measurement Methods: A Historical Overview" Sensors 22, no. 5: 1871. https://doi.org/10.3390/s22051871

[24] Szilágyi Zoltán, Hajdu Csaba, Bájári Patrik. Identification of position by means of shape recognition in mixed reality systems. In: Petőné Csuka, Ildikó; Simon, Gyula (szerk.) AIS 2023 - 18th International Symposium on Applied Informatics and Related Areas – Proceedings. Székesfehérvár, Magyarország : Óbudai Egyetem (2023) 177 p. pp. 80-84. , 5 p.

[25] Zoltán Szilágyi, Csaba Hajdu. Design and Requirement Analysis of an Outdoor Autonomous Mobile Manipulator. In: Petőné Csuka, Ildikó; Simon, Gyula (szerk.) AIS 2022 - 17th International Symposium on Applied Informatics and Related Areas – Proceedings. Székesfehérvár, Magyarország : Óbudai Egyetem (2022) 180 p. pp. 44-49. , 6 p.

[26] Hajdu Csaba, Boros Norbert, Gömböcz Zoltán, Szilágyi Zoltán, Ballagi Áron. Reactive Human-Robot Interaction in Digital Reality. In: IEEE International Conference on Cognitive Infocommunications (szerk.) 12th IEEE International Conference on Cognitive Infocommunications (CogInfoCom 2021) : proceedings. Online kiadás, Nemzetközi : IEEE (2021) 1,098 p. pp. 949-950. , 2 p.

[27] T., Jancso ; A., Podor ; E., Nagyne Hajnal ; P., Udvardy ; G., Nagy ; A., Varga ; Q., Meng ; L., Zhang. Data Integration with Geographic Information System Tools for Rural Environmental Monitoring. In: IRC 2022 XVI. International Research Conference Proceedings. Peking, Kína (2022) 270 p. pp. 52-59. , 8 p.

[28] Peter, Udvardy; Levente, Dimen; Gergely, Vakulya. Data Collection In The Wild: Challenges And Solutions. University of Agronomic Sciences and Veterinary Medicine of Bucharest. Scientific Papers. Series E. Land Reclamation, Earth Observation And Surveying, Environmental Engineering 12 pp. 167-175. , 9 p. (2023)

Resistance Spot Welding Control from Recycled Components

Csaba Nagy
Obuda University
Alba Regia Faculty
Székesfehérvár, Hungary
sabaka241@gmail.com

Bertalan Beszédes
Obuda University
Alba Regia Faculty
Székesfehérvár, Hungary
beszedes.bertalan@uni-obuda.hu
https://orcid.org/0000-0002-9350-1802

Abstract - This article provides a comprehensive overview of microcontroller-controlled resistance spot welding machines, emphasizing both the design and operational aspects, as well as the benefits of using recyclable components. It details the electrical and mechanical parameters of the equipment, along with the control electronics and regulatory processes. Additionally, the physical principles of resistance welding are explained, followed by an analysis of the main challenges in control theory related to the process. Solutions to these issues are presented, focusing on the types of measured signals and the control strategies employed.

Keywords - spot welder, resistance spot welder, welding current control, recyclable electronics, reusable electronics

I. INTRODUCTION

Resistance spot welding is one of the most common welding processes in industry, especially in the automotive industry and other metalworking fields. During the process, the workpieces are held between electrodes, where heat is generated as a result of a high current, which melts the surface of the metal and creates a permanent bond with the help of pressure. The success of the process is highly dependent on the precise control of welding parameters such as current, duration and electrode force, which is a challenge due to changing environmental conditions and different materials used [1-4].

This study aims to provide a comprehensive picture of modern approaches to the control of resistance spot welding, with particular attention to the application of microcontrollers and the integration of recyclable components. Microcontrollers enable accurate and fast control of welding parameters, while the use of recycled components reduces costs and enables more sustainable production processes. The purpose of the thesis is also to reveal the physical background of the welding process, problems related to control, and different measurement and control strategies.

Resistance welding is a welding process that uses heat generated by electrical resistance between workpieces to fuse metals together. In the process, two electrodes are placed between the workpieces, through which a high-current electric current flows, while the electrodes apply pressure to the metals. The increase in temperature results in local melting on the surface of the metals, which creates the welding joint with the help of the electrode force. The use of spot welding is particularly widespread in mass production because it is fast, can be automated, and does not require any added material such as shielding gases or welding wire [5,6].

II. THE PHYSICS OF THE SPOT WELDING PROCESS

In resistance welding, electric current flows through the electrodes through the workpieces, and the electrical resistance between the materials generates heat, which melts the surface of the metal. Heat production is based on the following equation:

$$Q_G = \int_0^{t_w} i_w^2(t) \sum_{i=1}^{n} R_i(t) dt \qquad (1)$$

where: Q_G is the amount of heat generated, i_w is the welding current, R_i is the resistive components present in the welding current path, t_w is the welding time and n is the number of measurements.

The amount of heat generated is directly proportional to the welding current and welding time. The largest part of the resistance is the contact resistance of the workpieces, which is crucial for the quality of welding. The resistance between the electrodes and the workpiece (approximately one-sixth of the previous value - for a steel workpiece and a copper electrode) is also significant, so these surfaces must be cooled frequently to avoid excessive heat generation and electrode wear.

A. The application of microcontrollers in the control of resistance spot welding

One of the biggest challenges in controlling resistance spot welding is the precise control and adjustment of welding parameters. Conventional welding machines operated with mechanical switches or simple timers, which did not allow dynamic control of the welding process. In recent years, the use of microcontrollers has brought significant progress in this field, as these devices allow accurate real-time control of welding current, voltage, time, and temperature [7].

The advantage of microcontrollers is that they have both digital and analog inputs, so they can process signals from various sensors, such as current, temperature, and pressure sensors [8]. Based on the sensed data, the controller is able to modify the welding process in real time to ensure the desired quality of the joint. Below is an overview of some of the key control tasks that can be performed efficiently with a microcontroller [9].

Continuous measurement and control of the welding current ensures that the workpieces receive the right amount of heat without overheating the material or insufficient heat generation. For this, a Rogowsky coil and a current sensor based on the Hall effect are often used, which enable fast and accurate measurement. Continuous monitoring of the temperature [10] during the welding process helps to avoid

overheating of the electrodes, which can lead to electrode wear and deterioration of welding quality. Based on the temperature measurement results, the controller can regulate the cooling system [11]. The quality and temperature of the welding can also be easily monitored with an acoustic microphone, by detecting the emitted infrared light, or by measuring the thermal voltage.

Setting the correct welding time is critical to the strength of the weld joint. The welding time can be accurately regulated with the help of a microcontroller depending on the thickness and type of the welded material. The advance of the electrodes, the applied pressure, is also controlled, and piezoelectric, ultrasonic or laser distance measurement can also be used.

B. Control strategies and solutions

The success of resistance spot welding is highly dependent on the choice and application of control strategies. The basic challenge of traditional resistance spot welding systems is that it is difficult to set and maintain welding parameters that ensure optimal welding strength, especially under changing conditions. Early RSW machines were controlled by mechanical switches and simple timers that allowed only basic parameter control. However, systems with programmed controls allow real-time monitoring and intervention of the welding process.

Modern resistance spot welding systems work with feedback control solutions that automatically adjust the welding parameters based on the measured signals. The control system uses sensors to measure welding parameters such as dynamic resistance, welding current or electrode displacement. Based on these, the output unit adjusts the current, voltage and welding time to ensure the desired joint quality.

The advantage of the feedback control system is that disturbances occurring during the welding process (such as changes in the surface condition of the workpieces, wear of the electrodes or fluctuations in the thickness of the workpiece) can be automatically corrected based on real-time data. In the following, the three main types of feedback control systems are presented: the direct feedback control system (FCS), the weld-to-weld control system (WWCS), and the monitoring system (MS).

The FCS system controls welding parameters in real time based on measured variables such as dynamic resistance or current. The essence of FCS is that the controller constantly compares the desired values with the actual measured data and intervenes in the process based on them – similar data processing tasks are presented in the following works [12, 13]. For example, if the welding current or dynamic resistance deviates from the desired level, the system automatically adjusts the current or welding time. This type of control is especially useful if there are frequent disturbances in the process, as it is able to correct them immediately. The advantage of the FCS system is that it ensures uniform welding quality even in the presence of disturbances during the welding process, thus ensuring the optimal quality of the welding joint. These algorithms and their advantages and disadvantages are detailed below.

Constant current control (CCC) is based on the assumption that the welding current must be kept stable during the welding process (1). The continuous measurement and control of the current makes it possible to ensure a constant value of the amount of heat during the welding cycle. The value of the current is determined in combination with the current transformer and hall sensor, thus also taking into account the contact resistance between the workpieces. The advantage is that keeping the current stable helps to accurately control the size of the welding nugget and the strength of the bond and is suitable for welding applications where sudden changes in the current are undesirable. The disadvantage is that the constant current algorithm does not take into account the changing dynamic resistance, so heat production may not remain optimal. In reality, the resistance often changes during the welding cycle, which can affect the welding quality, and is not applicable to a wide range of materials, where different welding parameters are required for different materials.

Constant voltage control (CVC) is similar to the constant current algorithm, but focuses on stabilizing the welding voltage (2). The continuous measurement of the voltage allows the welding current and heat production to remain at the correct level. The algorithm focuses on keeping the welding voltage at a constant level, thus reducing the effect of current fluctuations during the welding cycle. The voltage is measured directly on the electrodes, so the quality of the welding process can be monitored. The advantage is that maintaining a constant voltage helps to stabilize the welding quality and ensure the optimal amount of heat, and measuring the voltage is often simpler than measuring the current. The disadvantage is that the algorithm does not take into account the changing dynamic resistance, which can also change in connection with temperature changes during the welding process. Furthermore, voltage control can be problematic when the resistance between the workpieces changes suddenly, since the voltage does not always reflect the actual welding condition.

$$Q_G = \int_0^{t_w} \frac{u_w(t)^2}{\sum_{i=1}^{n} R_i(t)} dt \quad (2)$$

where, u_w is the voltage measured on the welding tips.

Constant power control (CPC) is the most advanced FCS approach that focuses on keeping welding power stable throughout the welding cycle. The algorithm regulates the welding performance by continuously monitoring the current and voltage, so that the optimal amount of heat remains at a constant level (3). This approach minimizes the risk of weld nugget size and temperature fluctuations. The advantage is that maintaining constant power reduces the risk of welding errors and ensures the optimal size of the welding nugget. It enables the welding of a wide range of different materials and thicknesses as it adapts to different material characteristics. The disadvantage is that the sensors and measurement technology required to stabilize the performance are more expensive, and the implementation of the algorithm is more complicated and requires more computing power from the control unit.

$$Q_G = \int_0^{t_w} u_w i_w dt \quad (2)$$

The tracking controller is designed to track a predetermined reference signal of different measured

variables, for example welding current, electrode movement, aquistic transmission [14-16]. PID control algorithms are often used for this purpose [17, 18].

Neural networks, fuzzy logic, and expert systems are the AI techniques commonly employed in resistance spot welding (RSW) applications. Expert systems are often used for the selection of appropriate welding parameters. Fuzzy logic-based control systems are typically used in tracking-type controllers, where dynamic resistance or electrode displacement is used as a reference. Neural networks are the most widely used AI techniques in RSW. They are employed for predicting weld nugget diameter, weld strength, or for weld classification. A commonly used network is the multilayer perceptron, which is combined with a backpropagation learning algorithm for function approximation; however, other types of neural networks have also been tested in RSW control [19, 20].

C. *Educational acpects and practical application of recyclable components*

The spot welding device, constructed from recycled components and described in this article, serves as an effective educational tool for demonstrating the control methodologies discussed in the previous subsection. This device is particularly valuable in higher education and engineering training, especially in fields related to sensors and actuators. It supports hands-on, hardware-focused applications, including microcontroller and industrial computer programming, as well as the teaching of algorithm theory and control systems technology [21-24].

One of the key aspects of modern industrial production is sustainability, which includes the efficient use of resources and the recycling of materials. The use of recyclable electrical components in resistance spot welding offers significant cost savings and environmental benefits. Integrating recyclable components not only reduces production costs, but also minimizes production waste, contributing to more sustainable industrial practices [25-28].

The benefits of recycled parts include cost reduction as companies need to purchase fewer new parts, thereby increasing their competitiveness. In addition, the use of new raw materials and energy costs can be reduced, as the incorporation of recycled components requires less energy than the production of new components. Moreover, recycling also contributes to job creation and innovation, as special labor is needed to disassemble and process used devices.

However, the use of recyclable components can present challenges, for example due to variability in quality and reliability. Certain recycled components may be outdated or no longer meet current technical requirements, which may reduce equipment performance and reliability. Therefore, when using such components, careful monitoring and testing is required, which also requires additional resource expenditure.

III. HARDWARE AND SOFTWARE IMPLEMENTATION

The transformer used is a 1500VA one. The outer winding is 400V and the inner winding is 230V. Then the outer winding had to be dismantled, the secondary winding was kept in its original state. After idle and load measurements, the following data can be measured: idle voltages U1=236V, U2=4.7V; load currents I1=17.8A, I2=889A, where U1 and I1 are the primary electrical parameters, U2 and I2 are the secondary electrical parameters of the transformer. The gear ratio is calculated as a=U1/U2≈50.

The frame had been built from aluminum profile from out-of-plant production machines, the construction can be seen on Fig. 1. The size of the workpiece and welding technology gives the design of electrodes. For the electrodes, the available 8 mm diameter piece of bronze was used. Textile bakelite was used to hold the electrodes, it insulates and heat-resistant, also tolerates mechanical stresses. With ten continuous welding, with only air cooling a 40°C temperature rise can be measured, see on Fig. 2.

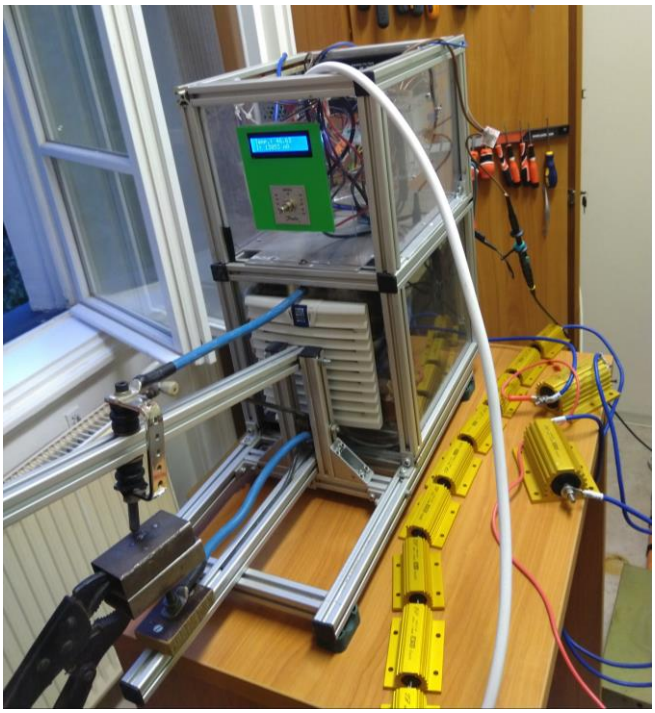
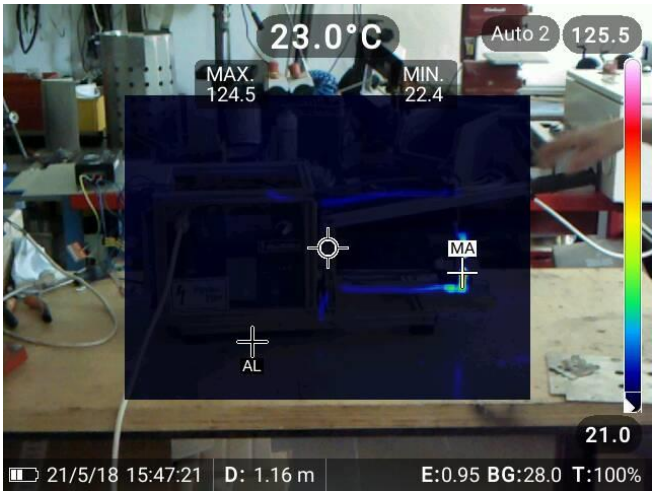


Figure 1: The finished equipment



2. Figure Thermal image of the electrodes after multiple welds

A. *Software & Control*

A microcontroller has been selected to control the device. A block diagram of the equipment can be found in Fig. 3. For user interface, a 16x2 line LCD display has been installed.

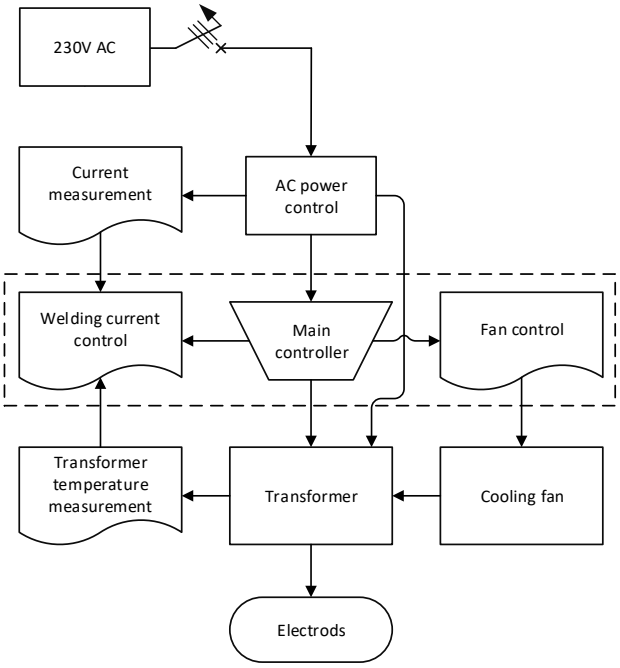


Figure 3: System design

Among several options for current measurement, a hall sensor current measuring module was selected and installed on the primary side of the transformer, for voltage measurement, an optocoupler based circuit been made to separate line voltage from the microcontroller, for temperature measurement, a silicon based component was selected.

Frequent use of equipment containing transformers creates a noise in the electrical network (Fig. 4. and Fig. 5), in order to reduce the effect, a properly designed mains filter has been installed and so that the disturbance at startup is acceptable, see on Fig. 6.

The transformer is continuously pre-excited through a resistor after sensing the welding tip contacts, when welding starts and the current increases, then the SSR2 (solid-state relay) will shunt out R1 and the current will flow through the relay. Timing is controlled from software, the process can be followed in Fig. 7.

Detailed description of the operation can be follow at Fig. 7. At start-up, it is checking is it the first start, and with electrical measurement it preventing the device from starting with burnt-in electrodes, in case of failure, it turns off the SSR1 relay and writes an error message to the display.

Temperature measurements are also performed to avoid overheating. If the preset value is exceeded, it turns off the equipment and informs the user with an error message.

Cyclically, the number of welds are counted. The electrodes can deteriorate during high-current welds. After the precalculated cycles, the operator receives a notification to check and clean the electrodes and then shuts off the device for the time of the operation.

After turning on the SSR1 relay, a current measurement is made, and if its value exceeds 2A, the SSR2 relay is turned on for the welding period. See the system architecture in Fig. 8.

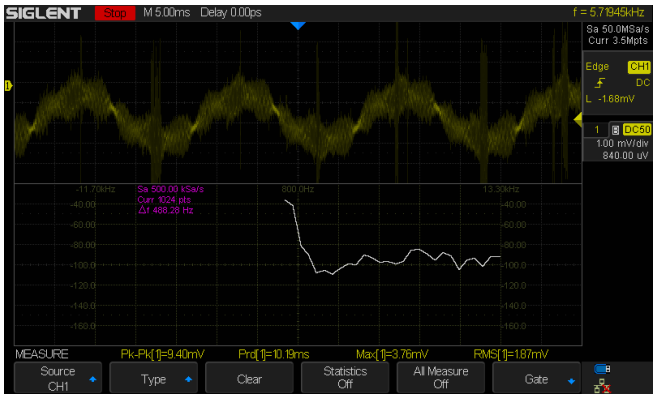


Figure 4: Noise spectrum before filtering

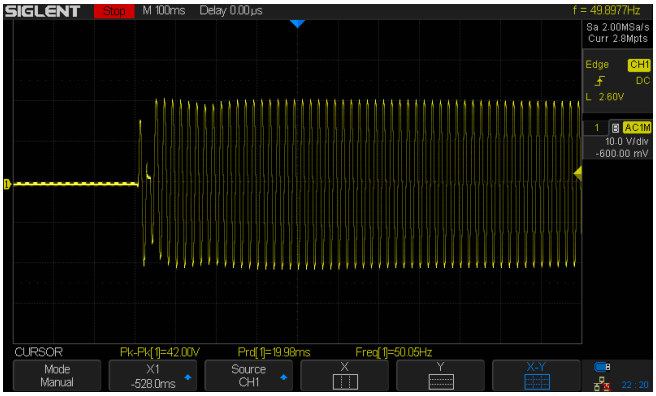


Figure 5: Noisy switching current diagram

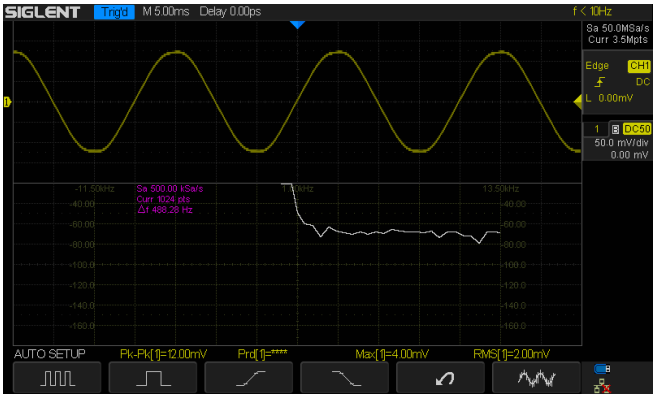


Figure 6: Noise spectrum after filtering

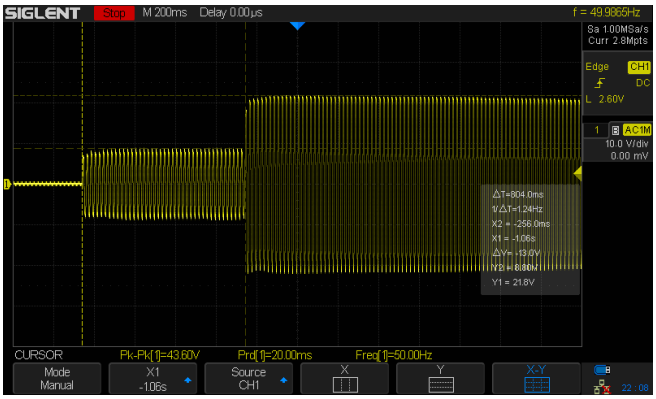


Figure 7: Timed welding process

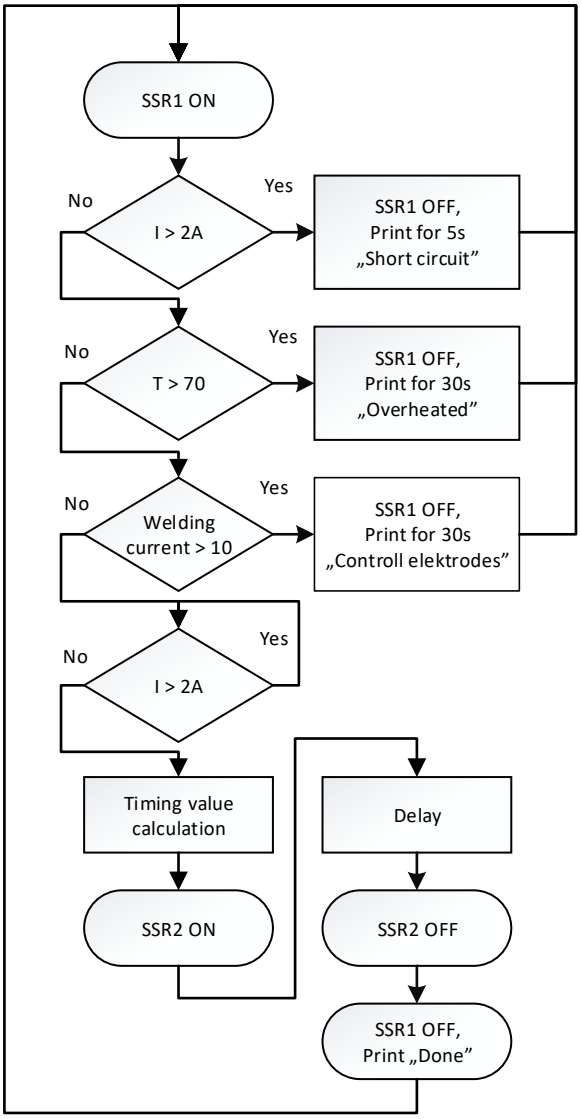


Figure 8: Flowchart of the main program

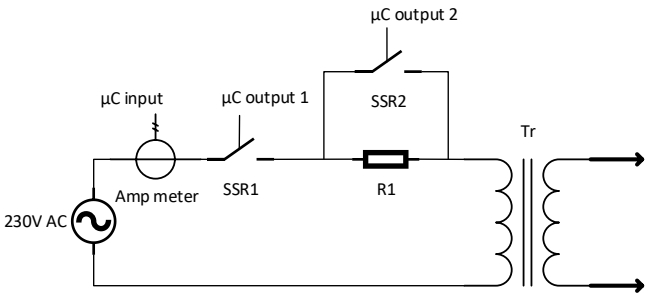


Figure 9: Primary side control elements

CONCLUSION

Resistance spot welding remains an essential industrial process, having undergone significant advancements with the integration of microcontrollers. These devices enable precise, real-time control of welding parameters, while the use of recycled components contributes to more sustainable manufacturing practices. Feedback control systems and monitoring of key variables in RSW play a critical role in improving weld quality and mitigating disturbances. Continuous innovation is vital for the welding industry to keep pace with technological advancements and address environmental and economic challenges. Additionally, the integration of control systems, microcontrollers, and recycled elements is not only crucial for the industry but also central to modernizing engineering education. By combining theoretical and practical learning, students gain valuable, industry-relevant knowledge in areas such as electrical engineering, materials science, and control systems, equipping them to become engineers who prioritize sustainability and innovation.

ACKNOWLEDGMENT

The authors would like to thank all the faculty staff and member that provide help and assistance throughout the project completion.

REFERENCES

[1] Overview of resistance spot welding control.P. Podržaj, I. Polajnar, J. Diaci, and Z. Kariž. Science and Technology of Welding and Joining 2008 13:3, 215-224

[2] Han, Z., Orozco, J., Indacochea, J. E., & Chen, C. H. (1989). Resistance spot welding: a heat transfer study. *Welding journal*, *68*(9), 363s-371s.

[3] Pouranvari, M., & Marashi, S. P. H. (2013). Critical review of automotive steels spot welding: process, structure and properties. Science and Technology of Welding and Joining, 18(5), 361–403. https://doi.org/10.1179/1362171813Y.0000000120

[4] Shen, Z., Ding, Y., & Gerlich, A. P. (2019). Advances in friction stir spot welding. Critical Reviews in Solid State and Materials Sciences, 45(6), 457–534. https://doi.org/10.1080/10408436.2019.1671799

[5] Kang Zhou, Ping Yao. Overview of recent advances of process analysis and quality control in resistance spot welding. Mechanical Systems and Signal Processing. Volume 124. 2019. Pages 170-198. ISSN 0888-3270. https://doi.org/10.1016/j.ymssp.2019.01.041.

[6] S. Aslanlar, A. Ogur, U. Ozsarac, E. Ilhan. Welding time effect on mechanical properties of automotive sheets in electrical resistance spot welding. Materials & Design. Volume 29, Issue 7. 2008. Pages 1427-1431. ISSN 0261-3069. https://doi.org/10.1016/j.matdes.2007.09.004.

[7] S. Ardabili, B. Beszedes, L. Nadai, K. Szell, A. Mosavi and F. Imre, "Comparative Analysis of Single and Hybrid Neuro-Fuzzy-Based Models for an Industrial Heating Ventilation and Air Conditioning Control System," 2020 RIVF International Conference on Computing and Communication Technologies (RIVF), Ho Chi Minh City, Vietnam, 2020, pp. 1-6, doi: 10.1109/RIVF48685.2020.9140753.

[8] Alexander Baklanov, Svetlana Grigoryeva, György Györök. Control of LED Lighting Equipment with Robustness Elements, Acta Polytechnica Hungarica (1785-8860 1785-8860): 15 3 pp 105-119 (2016)

[9] G. Gyorok and M. Mako, "Configuration of EEG input-unit by electric circuit evolution," 2005 IEEE International Conference on Intelligent Engineering Systems, 2005. INES '05., Spain, 2005, pp. 255-258, doi: 10.1109/INES.2005.1555168.

[10] Thermo-dynamic Cycle Computation of a Micro Turbojet Engine. Fozo, L; Andoga, R and Kovacs, R. 17th IEEE International Symposium on Computational Intelligence and Informatics (CINTI). 2016 17TH IEEE International Symposium on Computational Intelligence and Informatics (CINTI 2016) , pp.75-79

[11] Aizhan Zhaparova, Dimitry Titov, Alexander Y. Baklanov, Gyorgy Gyorok. Study of the Effectiveness of Switching-on LED Illumination Devices and the Use of Low Voltage System in Lighting. Acta Polytechnica Hungarica (1785-8860): 12 5 pp 71-80 (2015)

[12] Katona, János ; Czimber, Kornél ; Pődör, Andrea. Land Consolidation based on Cluster Analysis. ACTA POLYTECHNICA HUNGARICA 14 : 4 pp. 141-154. , 13 p. (2017)

[13] Katona, Janos ; Gulyas, Margit Horoszne. Determination of factors modifying land value based on spatial data. In: Drótos, Dániel; Vásárhelyi, József; Czap, László; Ivo, Petráš (szerk.) Proceedings of the 19th International Carpathian Control Conference (ICCC 2018).

Piscataway (NJ), Amerikai Egyesült Államok : Institute of Electrical and Electronics Engineers (IEEE) (2018) pp. 625-628. , 4 p.

[14] Mosavi Amirhosein, Bertalan Beszedes, Imre Felde, Nadai Laszlo, Gorji Nima E. Electrical characterization of CIGS thin-film solar cells by two- and four-wire probe technique. MODERN PHYSICS LETTERS B 2020 p. 2050102 , 16 p. (2020)

[15] Vakulya Gergely, Hajnal Éva, Udvardy Péter. Experimental Bolus Sensor for Dairy Cattle. 2022 IEEE 20th Jubilee International Symposium on Intelligent Systems and Informatics (SISY), 000157-000162, 2022, IEEE

[16] Chertov, A. M., and R. Gr Maev. "Determination of resistance spot weld quality in real time using reflected acoustic waves. Comparison with through-transmission mode." *16º WCNDT, Montreal* (2004).

[17] Song, Yuke, et al. "Research on the detection model and method for evaluating spot welding quality based on ultrasonic A-scan analysis." *Journal of Nondestructive Evaluation* 35 (2016): 1-12.

[18] Zhou, Kang, and Lilong Cai. "A nonlinear current control method for resistance spot welding." *IEEE/ASME Transactions on Mechatronics* 19.2 (2013): 559-569.

[19] Haefner, K., et al. "Real time adaptive spot welding control." (1991): 104-112.

[20] M. El-Banna, D. Filev and R. B. Chinnam, "Intelligent Constant Current Control for Resistance Spot Welding," *2006 IEEE International Conference on Fuzzy Systems*, Vancouver, BC, Canada, 2006, pp. 1570-1577, doi: 10.1109/FUZZY.2006.1681917.

[21] Kang Zhou, Ping Yao, Lilong Cai. Constant current vs. constant power control in AC resistance spot welding. Journal of Materials Processing Technology. Volume 223. 2015. Pages 299-304. ISSN 0924-0136. https://doi.org/10.1016/j.jmatprotec.2015.04.016.

[22] N., Tolner, M. Pogátsnik, "The large language models in the service of education", Magyar Nyelvőr, 148. Special Issue in English (2024-5) pp. 629-649. , 20 p., 2024.

[23] J. Módné Takács, M. Pogátsnik, "The Presence of Cybersecurity Competencies in the Engineering Education of Generation Z", Acta Polytechnica Hungarica 21, 6 pp. 107-127. , 21 p., 2024.

[24] J. Módné Takács, M. Pogátsnik, "The online learning from the students' perspective" In: Szakál, Anikó (szerk.) IEEE 19th World Symposium on Applied Machine Intelligence and Informatics (SAMI), Budapest, Magyarország : IEEE Hungary Section (2021) 507 p. pp. 00027-00032. , 6 p., 2021.

[25] Z. Trenka, V. Varsányiné Szeredi and M. Pogátsnik "The impact and potential of artificial intelligence in language education"in MAGYAR NYELVŐR 148 : Special Issue in English (2024-5) pp. 682-695. , 14 p. (2024)

[26] Biswajit Debnath, Priyankar Roychowdhury, Rayan Kundu. Electronic Components (EC) Reuse and Recycling – A New Approach towards WEEE Management. Procedia Environmental Sciences. Volume 35. 2016. Pages 656-668. ISSN 1878-0296. https://doi.org/10.1016/j.proenv.2016.07.060.

[27] Wang, H. (2002/09//). Using component knowledge base - a framework for building an electrical service system with reusable components. *Asia Pacific Management Review, 7*(3) Retrieved from https://www.proquest.com/scholarly-journals/using-component-knowledge-base-framework-building/docview/1115662723/se-2

[28] M. Tavakoli, P. AlhaisLopes, A. Hajalilou, A. F. Silva, M. ReisCarneiro, J. Carvalheiro, J. MarquesPereira, A. T. deAlmeida, 3R Electronics: Scalable Fabrication of Resilient, Repairable, and Recyclable Soft-Matter Electronics. Adv. Mater. 2022, 34, 2203266. https://doi.org/10.1002/adma.202203266

[29] Conti, Massimo, and Simone Orcioni. 2020. "Modeling of Failure Probability for Reliability and Component Reuse of Electric and Electronic Equipment" *Energies* 13, no. 11: 2843. https://doi.org/10.3390/en13112843

POSTER SESSION IV.

Education, Enterprise and Economics

Internal Training as a Retention Tool: Experiences from a Domestic Company

Ildikó Csuka Petőné
Alba Regia Faculty
Obuda University
Székesfehérvár, Hungary
csuka.ildiko@amk.uni-obuda.hu

Zsuzsanna Kaposvári
HR and Sales Department
Fejérvíz PLC
Székesfehérvár, Hungary
kaposvarizsuzsanna8106@gmail.com

***Abstract*— The aim of the study is to explore the role of internal corporate training in employee retention, focusing particularly on a domestic company where the training programs have recently started. In the problem statement the challenges of the labor market are presented, including the increase in turnover and the decline in employee commitment, which also emerge as part of global trends, as well as the corporate reasons behind this. During the literature review, authors analyze international and domestic sources, including various research findings that emphasize the effectiveness and economic return of training programs. Since the impact of the company's internal training programs are directly measure yet, the study focuses on predicting the initial effects of the training and the expected outcomes regarding employee commitment and development. The hypothesis is that the introduction of training may have a positive impact, which not only aids employee development but can also enhance the company's competitiveness.**

In the study, authors focus on analyzing the documentation of the initial training sessions already conducted by the company. This includes evaluating the training materials, participant feedback, as well as the goals and outcomes of the training programs. Additionally, a comparison will be made between the previous and current performance indicators of the company, such as turnover rates. This method allows for predicting the initial impacts of the training and understanding employees' expectations regarding their development opportunities. Furthermore, investing in human capital offers numerous benefits for the company: enhancing employee satisfaction and commitment reduces turnover, allowing the company to save significant recruitment and training costs. A skilled workforce can also bring innovative solutions that promote the company's competitiveness in the market..

In drawing conclusions, authors find that internal corporate training plays a fundamental role in employee development, increases employee satisfaction and commitment, and enhances performance, thereby potentially contributing to the long-term success of the company. At the end of the study, recommendations are formulated for optimizing future training strategies, emphasizing the importance of a culture of conthuman capital investment

Keywords - internal training, human capital investment, human resources turnover, workforce retention, employee commitment, competitiveness, human resources, labour market

I. INTRODUCTION

The Hungarian labour market is now unanimously characterised by a labour shortage. As a result, there is a growing preference for skilled and experienced workers already in the workplace. As it becomes increasingly difficult and costly for companies to recruit the right people from the labour market, they are raising wages and benefits, and are also seeking to benefit their existing employees by offering working conditions and career opportunities. Investing in human resources is now essential for companies, as it is their employees who are their greatest asset. A well-educated and motivated workforce not only increases productivity, but also a company's competitiveness. Ensuring the career path of employees is therefore essential for the success of companies operating in today's socio-economic environment. In order to ensure continuous development and to remain competitive, companies must actively support the professional development of their employees. Training not only increases employees' knowledge but also their motivation. Gaining new knowledge and skills helps career progression and helps employees keep up with the latest industry trends. The training provided by companies also offers opportunities to build professional contacts. The emphasis on teamwork and communication skills increases effectiveness. Company culture also improves when employees are constantly learning. Training makes companies' employees a more engaged workforce.

Companies should monitor the effectiveness of training programmes. Evaluations and feedback can help inform future improvements and help fine-tune programmes, which can lead to even greater satisfaction. Companies that continually update their training materials will become more attractive to talented employees. Internal training can be flexible, adapting to the needs of employees. This flexibility also helps to reduce turnover. Employees are more likely to stay with a company that offers opportunities for development. Integrating career planning and training strengthens employee engagement. Well-designed training pathways also help to develop future leaders. Developing leadership skills is key to the long-term success of companies.

For companies, training is an investment that pays off in higher productivity. Training gives employees a better understanding of the company's goals and values. The shared learning experience also strengthens team spirit.

Companies should offer different forms of training. Programmes tailored to individual needs make training more attractive. Developing soft skills is as important as technical knowledge. Enhancing creativity and problem-solving skills contributes to business innovation. Training also boosts workers' self-confidence.

To make the best use of resources, it is worth adapting training to the workflow. Some of the training opportunities may be non-formal, such as workshops or seminars.

Companies need to keep training materials up to date to remain relevant.

Internal training also promotes collaboration between teams. Learning together fosters a sense of community, which reduces turnover. Employees can build friendly relationships, which increases workplace loyalty. Training helps employees better understand the company's goals and values. This commitment to a common goal increases engagement.Internal training helps employees to integrate more easily into the company culture. A strong corporate culture is one of the cornerstones for reducing staff turnover. Employee identity becomes intertwined with the company, which increases loyalty. Internal training makes employees more confident in their jobs. Increased confidence reduces stress levels, which also contributes to employee satisfaction. More satisfied employees are less prone to turnover. Companies should therefore focus not only on knowledge but also on mental well-being. In-house training can therefore also lead to significant savings in the long term by reducing turnover.

The retention potential of in-house training in companies is illustrated from an international and a domestic perspective, and in more depth through a specific case study of a domestic company.In the study, the name of the company concerned is given anonymously.

II. International and National Perspectives on In-House Training

Let's first look at some international companies where the return on this training investment has been seen in the recent past.

Google has a reputation for investing heavily in employee welfare and training. The company offers a variety of programmes, such as 'Google Career Certificates', which help employees learn new skills. This approach has not only increased employee satisfaction, but has also boosted the company's innovation, contributing to high levels of product development.

Salesforce also pays significant attention to human resource development. The company spends millions annually on training and employee development. The "Ohana Culture" program emphasizes a sense of community and teamwork. As a result, employee turnover has been reduced and the company has been able to consistently increase revenues.

Accenture continuously invests in employee development and provides opportunities for learning and professional growth. The company has trained millions of people worldwide through its Skills to Succeed programme. As a result, the company has not only built a strong reputation in the labour market, but has also significantly increased the efficiency of its projects.

Let's look at some examples of companies in Hungary that have successfully invested in human resources and experienced positive returns as a result.

Richter Gedeon Pharma continuously invests in employee training and development. The company's internal training programmes, such as the "Richter Academy", provide employees with the opportunity to acquire new skills and advance in their careers. As a result, staff turnover has remained low and the company's capacity for innovation has increased.

Telenor actively invests in employee development, especially in digital skills. The company offers various programmes, such as the "Telenor Academy", which aim to enhance employees' knowledge and support career opportunities. As a result of these investments, Telenor has been able to maintain a stable workforce and increase employee satisfaction.

These examples illustrate that companies in Hungary also recognise the importance of investing in human resources. Training and development programmes have not only increased employee satisfaction but have also contributed to the success and long-term sustainability of companies.

Gary Becker's work [1] is related to human capital theory. In his book "Human Capital", Becker argues that training and education of the workforce increases the productivity of the individual, which ultimately benefits the firm. His theories have established the economic importance of investing in human capital. Becker and Huselid [2] have shown that well-trained workers contribute directly to improved firm performance. Their study suggests that investing in human capital leads to higher profitability in the long run. Research by Guest [4] showed that the relationship between workplace well-being and employee commitment is strong. Training and development at work makes employees feel better about themselves, which reduces turnover and increases productivity. Brewster and Collings [3] emphasise that it is particularly important for international companies to adapt to the needs of the local labour market. Tailoring training programmes improves employee satisfaction and workplace culture, which helps retain talent.

Finally, Kirkpatrick's [5] model is used to measure the effectiveness of training and emphasises that the success of training programmes depends not only on increasing the knowledge of participants but also on putting it into practice. According to the model, effective training contributes to improved employee performance and company results. Kirkpatrick's four-level training evaluation model gives companies a comprehensive picture of training effectiveness. The first level, reaction, measures how satisfied participants are with the training. This feedback is important because positive experiences encourage further learning. The second level, learning, assesses the extent of knowledge and skills acquired. This data allows companies to understand whether training really helps employees to develop. The third level, behaviour, monitors the use of applied knowledge and skills in the workplace. If workers are able to apply what they have learned, this has a positive impact on the company's performance. The fourth level, outcomes, measures the economic impact of training, for example through increased production, reduced costs or increased profits. The application of the Kirkpatrick model allows companies to compare their training costs with measurable results, so that it is clear why it is worth investing in training.

Kotter's work [6] on change management also emphasises the role of corporate training. Kotter argues that for change to be successful, it is essential that employees are prepared and motivated. Training helps employees better understand the need for change and their own role in the transformation. This engagement and understanding leads to a stronger organisational culture that helps the company to successfully implement change. Training develops not only technical skills but also soft skills such as communication, teamwork and problem solving. These skills are key to improving workplace

collaboration and managing conflict. Well-trained employees who can work together effectively contribute to the success of the company and improve the workplace climate.

Corporate training is therefore not only a cost, but also an investment in the future. Employee development reduces turnover because employees become more loyal to the company if they feel that the company cares about their development. In the long run, this loyalty reduces recruitment and training costs, so that training pays off economically.

Karoliny [7] stresses the importance for companies of achieving a work-life balance. A good working environment and flexible working increase employee satisfaction, which ultimately leads to improved productivity. The opportunity for continuous learning and development is essential to maintain motivation in the workplace. These principles apply to managers and employees alike. Managers must be conscious of their employees' development and support them in achieving their goals. “Each employee needs to have specific goals that contribute to and are aligned to the organisation’s strategic objectives, goals and vision.” [12]. Developing human resources therefore not only increases costs but also pays off in the long run.

With well-designed training and development opportunities, companies will be able to get the best out of their employees. Tailored training programmes can help employees to feel better about their jobs, which contributes to increased job satisfaction. In addition, training promotes cooperation between employees and improves team dynamics. Shared learning experiences strengthen the relationships between employees, which leads to an improved working atmosphere. The trust and cooperation developed during training contributes to the success of the company. Karoliny points out that investing in human resources also contributes to a stronger corporate culture. And a positive corporate culture increases employee loyalty and satisfaction. According to Ildikó Petőné Csuka [8], "in order to be competitive, organization should organise training for their employees ... that motivates them and makes them committed to the organization".

According to the researchers, knowledge sharing and collaboration are the basis of the workplace of the future. Trust and respect between employees fosters creative thinking that can lead to new solutions. Investing in human resources is therefore essential not only for the present but also for the future.

This data confirms that investing in company training has a positive impact not only on employee development but also on company performance. Poór's research provides extensive evidence of the effectiveness and economic return on training, reinforcing the importance of human resource development.

In many cases, training structures are far from being in line with the interests of the business and this creates a gap between the two sides. The shortages in the quantity and quality of labour in some sectors can be attributed to two main factors. The first is the so-called "skills gap", where adults of working age lack the right skills, often not even the basic competences needed. The other is the so-called "skill mismatch", where the structure of the economy's human resource needs does not match the skill structure of adults, i.e. the knowledge acquired through training does not match what is expected and required within the company. This can be addressed and reduced by trying to identify the two interests and prepare individuals according to the basic competences already at the basic education stage [11] .

In the company analysed in the present study, both occur, but the latter is a bigger problem, more precisely, the training of the workforce required by the company has been removed from the state training structure (due to lack of interest), i.e. the training of the specialists required by the company has been removed from the state framework, causing a big puzzle for F. Zrt. as to how to solve the problem of labour supply. The company had to take over the training of sectoral specialists itself, in cooperation with local vocational schools. In the following this problem will be outlined below, together with the responses to the problem, the strategies and tactical and operational training plans set, the implementation of the training plans, the assessment of their effectiveness, their evaluation and, in the light of this, future training development plans.

III. LABOUR MARKET CHALLENGES IN THE WATER UTILITIES SECTOR

“Several large companies recognized the situation that they could only ensure their labor supply continuously if they participated intensively in vocational training.” [10] In 2019, the Hungarian Water Utility Association published its impact study on the water utility sector, entitled "Summary of the labour shortage present in the water utility sector" [13]. In the study, the problem of supply and shortage of professionals was formulated, in the background of which the following causes were identified during their research work:

- education in the school system for special professions needed by the water utility sector has been visibly reduced;
- due to the continuously decreasing number of students, educational institutions training skilled workers have been dismantled;
- the prestige of manual labor has also been declining;
- It was observed that due to the low number of students, interested persons were directed to other fields.

The problems affecting the sector also appeared in the everyday life of F. Zrt. The low wage level before 2023, the ageing age pyramid (“Fig. 2”) and the lack of supply caused significant fluctuations in the organization (“Fig. 1”). During 2023, the number of leavers was much higher than in previous years. It was then that the company decided that further human resources strategy measures were needed to retain and train its workforce.The involvement of recruitment and skilled workers has been hindered. In the 2018/2019 school year, according to the data provided by the Educational Authority, a total of 565 people studied professions belonging to water management professional groups in 12 vocational secondary schools for the entire territory of the country:

- Water technician 238,
- Water management technician 254 people,
- Water utility technician 35 people,
- Hydraulic engineering technician 5 people

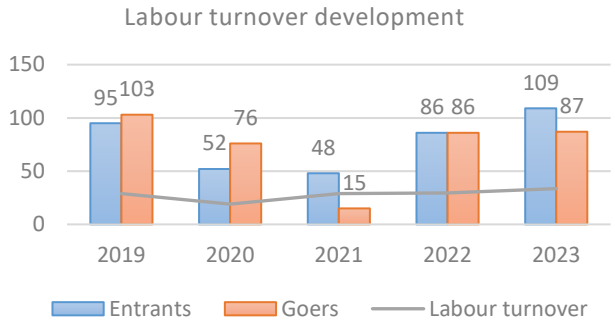


Fig. 1.: Labour turnover development [14]

For the company, this meant that it had to take new human resources strategy measures and formulate new guidelines in the field of education [14].

The ageing age pyramid also emphasizes the importance of youth education and training. (“Fig. 2”)

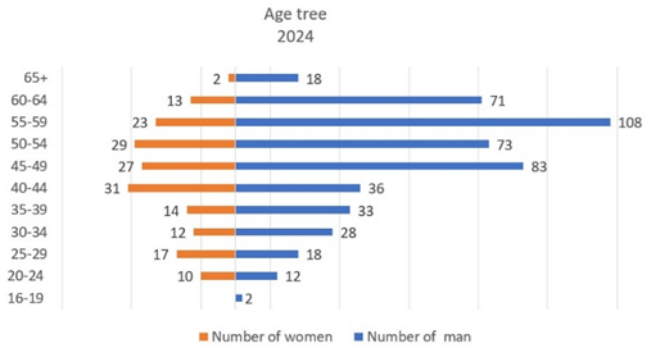


Fig 2.: Age tree [14]

A. New corporate human resources strategy

The following objectives were set:

1) The company's education system needs to be renewed
2) Internal training should be broadened and knowledge sharing should be facilitated
3) Educational institutions should be contacted
4) Join the dual training system
5) The individual development path of employees should be facilitated for training plans, and workers' training needs should be assessed. Individual and group training plans should be prepared
6) In sectors lacking labour, youth education must be prioritized
7) Further training of manual workers should be established in the vocational training of plumbers who are deficient in water and sewerage systems
8) A new training offer for white-collar workers should be implemented

“In a dual training system, the parties are linked together in a joint process of knowledge and experience development, which implies close, result-oriented cooperation between the partners” [9].

As a step towards achieving these goals, partnership cooperation was launched with the following institutions:

• Vörösmarty Mihály Technical and Vocational School – vocational training

• Székesfehérvár SZC Hunyadi Mátyás Technikum – vocational training

• Székesfehérvár Vocational Training Center – vocational training, dual training program

• Óbudai University – higher education, dual training program at the following faculties and trainings

o Alba Regia Faculty of Engineering
Bachelor of Surveyor, Landscaping Engineer
Bachelor of Engineering Information Technology
o Károly Keleti Faculty of Economics
Bachelor of Business Administration and Management
Technical Manager Bachelor /Budapest/
o Sándor Rejtő Faculty of Light Industry and Environmental Engineering

Technical Manager Bachelor

• Chamber of Commerce and Industry of Fejér County – promotion of the profession, awareness-raising, registered dual training centre status, training of practical instructors

Within the company, the needs of employees were revealed with the help of a questionnaire. The questionnaire was completed by 55 people, and the results revealed that 87% of the respondents would require training.

IV. TRAINING PRACTICE OF CLASSES

As a result of negotiations with the Vörösmarty Mihály Technical and Vocational School, the first "F. Zrt. class" could start in autumn 2022 in the Water and Sewerage Network Mechanic training. The following can be summarised about the gap-filling training:

• Training in the HVAC sector, which ends with the acquisition of a vocational qualification.

• Duration of training: 1 year

• Curriculum: The plumber and sewerage installer builds water, firewater, sewer and water technology networks from various pipe materials. It installs and installs individual and central fixtures, domestic hot water producers, heat exchangers, water engineering system components, taps and fittings. It reviews existing networks and systems, reveals errors and repairs them. It builds a water and sewerage network within the boundaries of the plot, installs meters, and prepares public utility connections under professional guidance. He carries out periodic inspection work related to his work, the necessary administrative activities, conducts handover tasks, keeps in touch with the customer.

• Competence requirement: Technical orientation, creativity, striving for independence, precision, problem-solving ability, working in teams, striving for quality work.

A special feature of the training is that the school started a separate class out of 14 students. During the training, workplace mentors helped the students. In the first semester of the training, students studied theoretical subjects taught by the school, such as electronic foundation, HVAC foundation, HVAC measurements, technical drawing knowledge or basic welding knowledge. The total number of hours was 1355 hours.

After the first semester, students took exams in theoretical subjects. All candidates of F. Zrt. passed the exam in January 2023. In the second semester there was a practical training that took place at our company. The practical subjects included water supply I, II, Sewerage I, II, and swimming pool technology.

It was important for the company to provide training not only in shortage professions, but also to create opportunities for the competence development of white-collar employees. After the success of the manual workers, a department was also started among white-collar workers, in cooperation with the Székesfehérvár SZC Hunyadi Mátyás Technikum.

The main features of the training:

• Training in the business and management sector, which ends with a high school diploma and the acquisition of a technician-level qualification.

• Curriculum/ output requirements: The qualified person is capable of performing the administrative tasks of the enterprise, participates in the performance of accounting and financial tasks, participates in the accounting and financial tasks of micro, small and medium-sized enterprises, and performs certain partial tasks independently. He keeps in touch with colleagues and customers and organizes events and meetings.

• Competence requirements: Agile, good communication skills, problem-solving and organizational skills, teamwork, sense of responsibility, openness to the use of IT tools.

The training lasted from September 2023 until May 2024. The total number of hours of training was 690 hours. Some of the classes took place online. Currently, 9 students are preparing for the final exam, which will take place in October 2024. The exam consists of three parts: (1) a practical accounting task, (2) a test series of theoretical material and (3) a defense of the portfolio.

Through practice and pre-exam preparation, we strengthen and develop more competencies of participants. Communication skills, professional knowledge, concentration skills all had to be used during practice. Thus, overall, an extremely complex development could be realized in this case as well.

A. Research survey results

In the specific feedback, participants indicated that their relationship with colleagues has become more positive, they have gained a lot of new knowledge, and they have gained much more practical knowledge than before. He got a better insight into knowledge, deeper knowledge. Closer collegial relationships were experienced as an advantage by almost all participants. They also said that during the training, colleagues became more persistent and active. Many considered the trainings complex and indicated the need for more basic knowledge.

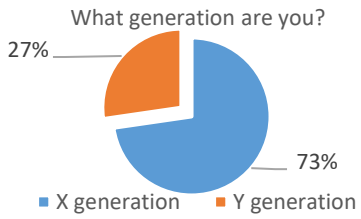


Fig. 3.: Generation distribution [14]

Three quarters of respondents are members of Generation X (1965-1980) (“Fig.3”).
Most participants participated in the training to increase knowledge, “TABLE 1.”

1. TABLE 1. [14]

What did the training give?	Distribution
A lot of new theoretical knowledge	72,7 %
Practical knowledge	9,1 %
Exam preparation	9,1 %
I already knew	9,1 %

After completing the training, more than 72% of participants did not lose loyalty to the company, while 27% did better. Although the survey shows that it is stagnant, the exit rate of participants is close to zero, meaning loyalty is reflected in actions related to the company. More than 90% of the participants applied for the training to increase their knowledge. The others considered it important to acquire knowledge because of a side hustle. More than 90% of the participants applied for the training to increase their knowledge. The others considered it important to acquire knowledge because of a side hustle.

The majority of participants (more than 36%) rated their knowledge expansion 8 on a scale of 10 on a scale of 10 through training. “Fig 4.”

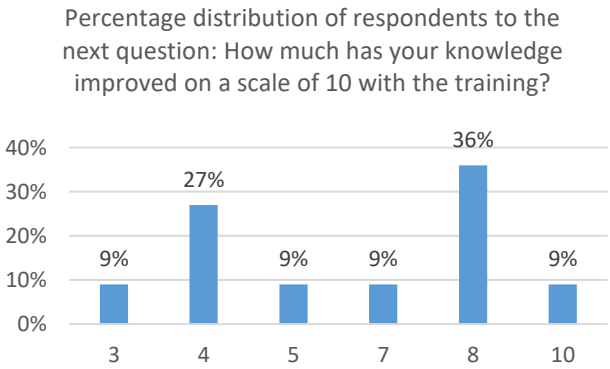


Fig. 4.: Knowledge scale [14]

B. Effects of training

- competency growth

In the months before the exam, students also practiced the exam task and completed the theoretical tasks. For the practical task, a practice board was set up in the workshop of the company's headquarters, on which the students could carry out the installation work. This developed their practical knowledge and technical skills.

For theoretical tasks, employees developed digital competences on laptops, since the majority of manual workers do not use such a device in their everyday lives. Their personal development throughout the process was greatly aided by these practices. There were participants who first encountered the development of digital skills here.
During the year, participants had to prepare five portfolios of tasks arising during their daily work. In addition to the text description, the portfolios illustrated the various cases with photos. Before the exams, these portfolios were practiced verbally with the help of mentors, which improved the communication skills of the participants.
During the exercises, the participants had to use speed, concentration and technical skills. The progress was visible, the entire class passed the exam in the profession of plumber and sewerage installer.

• knowledge sharing

Knowledge sharing was an important goal during the trainings. We talked about participants working in different fields in both mental and physical staff, who shared their knowledge with each other through the trainings. Water network workers went over to the sewer system to practice and vice versa. During the training, white-collar employees helped each other by exploiting competences, and the partner departments shared their theoretical experiences with each other. Based on the feedback, this is a cardinal point of the system, as it is important not only for knowledge sharing, but also for the partnership that increases performance in all areas.

• reduction in fluctuation

Partly as a result of the increase in education and training, the turnover of labour at the company decreased. For 2024, half of the previous year, inflows and outflows at the company decreased. This means greater stability. The trainees concluded a training contract, which contributed to the stability of the workforce. “Fig 5.”

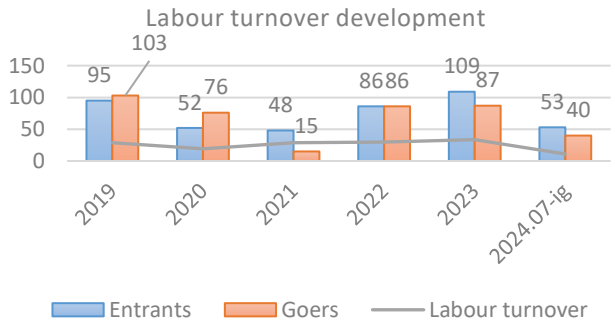


Fig. 5. : Labour turnover develop [14]

• increase in the number of participants in various trainings

The launch and success of the own department has increased the appetite for training among employees, who applied for training courses in 2024 twice as many as in the previous year. It was noticeable that the success of the first training had an incentive for many to complete new trainings. “Fig 6.”

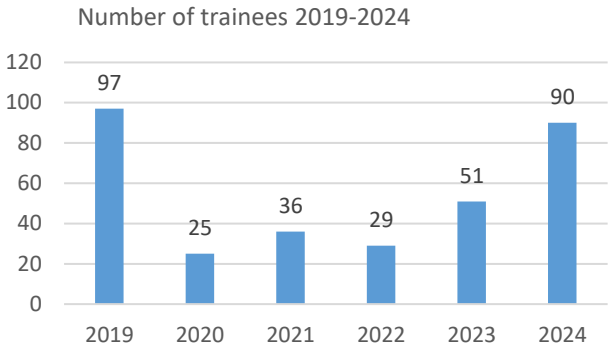


Fig. 6.: Number of trainees [14]

V. PROPOSALS FOR OPTIMISING FUTURE TRAINING STRATEGIES

The future training strategy is fundamentally based on the needs of the company adapted to the processes. The issue of youth education, the monitoring and replacement of shortage professions is decisive in the corporate training palette. Consideration of employee needs remains inevitable. The importance of corporate learning lies in the fact that the employee lays the foundation for a qualitative increase in the company's human capital by increasing his own competence values.

A. Based-on feedback, future trainings should include more practical motives, strive to further broaden competences
B. Trainings should be tailored to the company, the necessary knowledge should be adapted to the work
C. Communicate expectations and present the syllabus with the applicants for the training
D. Further develop the corporate learning culture, highlighting the importance of knowledge expansion and competence
E. More opportunities to practice should be provided during trainings.

It can be concluded that the training is definitely useful for the company, if the participants show interest in the subject of the training, the desire to learn will also multiply many times. Planting a learning culture sets an example for employees, and developing culture in this direction increases the desire to learn. One of the main advantages is the connection of participants in a training, who later help each other in everyday work, so the improvement of employee relationships has a positive effect on performance. The increase in competencies basically brings knowledge expansion and quality development, which the company can later build on, apply job broadenings and increase performance. The key is employee interest in training and active participation, for which it is advisable to launch trainings that are as tailored to the needs as possible. The knowledge expansion of employees clearly increases the company's profits, improves the qualitative composition of human resources and has a positive impact on performance through better employee relations.

VI. Summary

The example of the domestic company (F. Zrt.) has shown that internal training plays a key role in reducing company turnover. New skills and knowledge acquired through training provide opportunities for career progression. Employees who can see their potential for development are more likely to stay with the company, i.e. they are more satisfied and this increases their commitment to the company. Regular training increases employees' professional competencies, which enhances performance. Companies thus gain valuable long-term employees. In our study, we have listed some international and Hungarian company examples before analysing the circumstances of training implementation and its effectiveness in a specific domestic company through empirical data.

Overall, the international and domestic literature, as well as the survey and results of the domestic company, confirm that investing in human resources is beneficial not only for employees but also for companies. Employee satisfaction, reducing turnover and increasing productivity are all factors that contribute to the success of companies in today's competitive market. Training programmes and workplace development are therefore not only a trend but also a strategic choice for companies operating in today's economic environment. The results of the company analysed in the study (F. Zrt.) show that investing in training is not only a cost but a future gain. Investing in the right training helps to increase employee loyalty and commitment. The sharp drop in turnover figures, the significant satisfaction with training, the image of development opportunities clearly made employees more committed to the company and had a team-building effect. Well-trained employees not only improve company performance, but also shape the organisational culture. Training not only helps employees to develop professionally, but also helps to improve company culture and team dynamics. The development of human resources therefore contributes to the long-term success of the company. Investing in training helps companies to focus not only on short-term goals but also on long-term strategies.

References

[1] Becker, G. S. (1964). Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education. University of Chicago Press.

[2] Becker, B. E., & Huselid, M. A. (1998). High Performance Work Systems and Firm Performance: A Synthesis of Research and Managerial Implications. Research in Personnel and Human Resources Management, 16, 53-101.

[3] Brewster, C., & Collings, D. (2004). The Internationalisation of Human Resource Management. In: Scullion, H. & Collings, D. G. (Eds.), Globalizing Human Resource Management (pp. 45-68). Routledge.

[4] Guest, D. E. (1997). Human Resource Management and Performance: A Systematic Review of the Evidence. International Journal of Human Resource Management, 8(3), 263-276.

[5] Kirkpatrick, D. L. (1994). Evaluating Training Programs: The Four Levels. Berrett-Koehler Publishers.

[6] Kotter, J. P. (2002). The Heart of Change: Real-Life Stories of How People Change Their Organizations. Harvard Business Review Press.

[7] László Gy., Poór J., Farkas F., Karoliny M. (2010): Human Resource Management Handbook, Complex Publishing

[8] Petőné Csuka, I. (2014): Human Capital Analysis among Higher Education Teachers in Nagy IZ (ed.): Business Development in the 21st Century, Volume IV (p. 301-310) Budapest, Obuda University Keleti Károly Faculty of Economics

[9] M. Pogatsnik, "Dual Training in Engineering Education," in IEEE 21st World Symposium on Applied Machine Intelligence and Informatics (SAMI 2023) Proceedings, A. Szakál, Ed., Herlany, Slovakia: IEEE, 2023, pp. 169-174.

[10] M. Pogátsnik, "Dual education: connecting education and the labor market," Opus et Educatio: Munka és Nevelés, vol. 8, no. 3, pp. 304-313, 2021.

[11] Sebők, M. (2017): Empirical research on labour shortages and labour mobility. In Knowledge Management. Periodical of the Faculty of Culture, Pedagogy and Rural Development of the University of Pécs Vol. XVIII No. 2 , 55-64.

[12] Veresné, Valentinyi K. ; Czeglédi, Cs.: Group Coaching in Service of Talent Management (2018) Conference presentation:III Carpathian Basin Education Conference June 22-23. 2018., Nagyvárad

[13] https://maviz.hu/assets/GINOP-5.3.5-Kornyezet--es-hatastanulmany.pdf

[14] Internal documents, research, studies of the F. Zrt.

Challenges and solutions in online examination

Nikoletta Tolner
Alba Regia Faculty
Obuda University
Szekesfehervar, Hungary
tolner.nikoletta@amk.uni-obuda.hu

András Dávid
Alba Regia Faculty
Obuda University
Szekesfehervar, Hungary
david.andras@amk.uni-obuda.hu

Monika Pogatsnik
Alba Regia Faculty
Obuda University
Szekesfehervar, Hungary
pogatsnik.monika@amk.uni-obuda.hu

Abstract: **With the digitalization of education, online examination has become an increasingly widespread method, not only for language tests but also in the examination processes of other educational fields. The aim of our research is to provide a comprehensive overview of the administration of online exams in various educational institutions, and to showcase the methods and procedures applied through practical examples. In our secondary research, we analyzed the online examination systems of several language test centers, as well as universities and other educational institutions. Special attention was given to the technical requirements, identity verification, and anti-cheating measures. Based on a review of the literature, we explored how the online environment affects test-taker performance and the credibility of the evaluation process. The results of our research highlight that, while online exams offer significant flexibility and broader accessibility, they also present serious technical and ethical challenges. Through practical examples, we illustrate how different educational institutions are utilizing these technologies and what steps are necessary for future developments to ensure the credibility and fairness of exams.**

Keywords: online examination, digital education, cheating prevention, exam credibility, academic integrity, online test security

I. INTRODUCTION

The emergence of online exams is closely linked to the digitalization of education, which began in the early 21st century. The broader availability of the internet, the rise of e-learning platforms, and technological advancements have all contributed to the gradual transformation of educational systems [1,2,5]. However, online exams became widespread primarily due to the impact of the COVID-19 pandemic, when traditional educational institutions were forced to move both teaching and examinations to remote platforms [3,7].

During the pandemic, the restrictions and school closures led to a rapid increase in demand for online education [4], which also brought the necessity for online examinations [6]. Technological tools like Zoom, Google Meet, and various e-learning platforms (e.g., Moodle, Canvas) became fundamental tools in the daily processes of learning and examination [7,8]. The introduction of online exams provided the opportunity for students to take exams in a safe environment from home without interrupting their studies [9, 10].

Online exams offer flexibility, as they make it more convenient for students to choose the time and location for taking their exams [9]. However, several challenges arose, such as preventing cheating [11], managing technical issues, and ensuring reliable identity verification. Educational institutions and exam centers have attempted to address these challenges with various technological solutions, including online proctoring systems, camera monitoring, and AI-based algorithms for exam surveillance [12,13,14].

Overall, COVID-19 played a key role in the rapid spread of online exams, and although many places have returned to traditional exam formats, online exams remain popular, especially in situations where flexibility and accessibility are crucial.

In the first part of this study, we will conduct secondary research to explore the advantages, disadvantages, and challenges of online exams. In the second part of the article, we will present examples of methods and techniques used to ensure the authenticity of exams. During the secondary research, we searched for scientific articles in the Scopus database and Google Scholar. The search term used was “online exam”.

II. ONLINE EXAM

Online exams can be conducted either using the examinee's own devices or the equipment provided by the educational institution. Both approaches have different advantages and disadvantages, such as access to technology, reliability, and the prevention of cheating (table 1). Later, we will explore the benefits of online exams in detail, such as flexibility and cost-effectiveness, as well as their challenges, including technical issues and reliable identity verification. Additionally, a separate section will address the phenomenon of cheating, and the preventive methods applied to combat it.

TABLE 1. COMPARISON OF ONLINE EXAMS: INSTITUTIONAL EQUIPMENT VS. STUDENT EQUIPMENT (OWN EDITING)

Criteria	Institutional Equipment (Advantages)	Institutional Equipment (Challenges)	Student Equipment (Advantages)	Student Equipment (Challenges)
Technical Reliability	Monitored, well-maintained devices and stable network	Students must travel to the institution	Students can use familiar devices	Varied devices and internet connections may cause issues
Access	Devices are available to everyone on-site	Location dependence may restrict students living far away	Can take the exam from any location	Some students may lack access to proper devices
Security	Physical supervision, easier to prevent cheating	Requires physical presence of all participants	More cost-effective, no travel required	Harder to prevent cheating or ensure supervision
Cost	Institution manages equipment and security	High maintenance costs	Lower cost for the institution	Students may need to invest in expensive devices
Student Experience	Uniform environment and technology	Stress from unfamiliar location and devices	Familiar environment reduces stress	Technical problems or distracting environments may affect performance
Institutional Control	Full control over environment and security measures	Time-consuming and limited flexibility	Flexible time and location for students	Harder to ensure proper supervision

A. Advantages of Online Exams

The greatest advantage of online exams is the flexibility they offer in terms of location and time. Unlike

traditional exams, which take place at a set time and location, online exams allow students to participate from anywhere with an internet connection. This flexibility is especially beneficial for non-traditional students, such as those working full-time, having family obligations, or living in different time zones. Moreover, online exams often adapt to students' preferences, enabling them to take the exam when they feel most prepared. This adaptability can reduce exam-related stress and alleviate the logistical challenges of traveling to exam centers [15,16].

Online exams significantly increase accessibility, providing opportunities for a broader range of students to participate. This includes international students who cannot access local exam centers, as well as students with disabilities or special needs, such as those entitled to extended time or assistive technology. Many online platforms are compatible with screen readers, speech recognition software, and other tools, making learning and examination more inclusive. This promotes educational equity by breaking down geographical and physical barriers that could otherwise prevent students from completing their studies [17,18].

The implementation of online exams can lead to significant cost reductions for educational institutions. Traditional exams require physical spaces, proctors, and printed materials, all of which incur costs. By transitioning to digital formats, institutions can save on these expenses. Additionally, students save on travel and accommodation costs, especially when they would otherwise need to travel long distances to reach exam centers. These savings can make education more accessible to a broader student community, further democratizing higher education [19]. Moreover, online exam platforms are easily scalable, allowing them to accommodate larger numbers of students without the need for additional physical infrastructure.

In conclusion, the main advantages of online exams are flexibility, increased accessibility, and cost reduction, which together make education more inclusive and efficient in the digital age. These benefits became particularly valued during the COVID-19 pandemic when online learning and exams became indispensable. Further refinement of these systems will allow them to become more seamlessly integrated into the broader educational environment.

B. Disadvantages and Challenges of Online Exams

The introduction of online exams brought several advantages, such as increased flexibility and accessibility, but it also raised significant challenges that both educational institutions and students face. One of the biggest challenges of online exams is dealing with technical difficulties [20]. The lack of a stable internet connection, issues with computers, and software incompatibility can cause significant stress for exam takers [21]. These problems not only negatively affect individual performance but can also call the exam's credibility into question. Providing appropriate technical support to all students, especially those in different locations and time zones, is a challenge for educational institutions.

Online exams handle sensitive data, such as students' personal information and exam results [21]. Adhering to data privacy regulations and ensuring the secure use of such data is of utmost importance, as technologies like video monitoring or online proctoring software are often used during the exam process [22]. However, these technologies raise privacy concerns, as they may access sensitive personal spaces of students, such as their homes. Finding the right balance between security and privacy protection is a key task for educational institutions.

One major challenge of online exams is the reliable verification of the exam taker's identity. While in traditional exams it is easy to verify in person that the right student is taking the exam, in an online environment the risk of cheating increases. Although various technological solutions exist for identity verification, such as biometric identification or online proctoring systems [23], they cannot fully guarantee the authenticity of the exam taker's participation.

The spread of online exams has also raised the issue of equal access [22]. Not all students have access to proper technological devices or a stable internet connection. Economically disadvantaged students or those living in remote regions face difficulties. It is a challenge for educational institutions to compensate for these inequalities, for example, through special support or alternative exam opportunities.

Taking exams in an online environment can decrease students' motivation and commitment to learning. Due to the lack of in-person presence, many students take online exams less seriously [21], and managing time and stress effectively can be a challenge. Additionally, without direct supervision from teachers, it may be easier for students to lose focus or procrastinate.

The assessment methods used in online exams often differ from those in traditional exams. Open-book tests, automated test grading, and other alternative assessment methods can be beneficial but also raise concerns about credibility [24]. It is a challenge for educators to evaluate students' performance objectively and fairly while ensuring the reliability of exam results [25, 26].

The technological advancement of online exams and the rise of digitization offer significant advantages but are not without challenges. Technical problems, data privacy, identity verification, unequal access, and motivational issues are all problems that educational institutions must address as online exams continue to spread. However, these challenges can be addressed with continuous development and proper support, ensuring that online exams remain an effective and equitable solution for students in the long term.

C. Cheating in Online Exams

One of the biggest challenges of online exams is the prevention and detection of cheating. While in traditional, in-person exams, invigilators can directly observe students' behavior, ensuring that all participants follow the rules is more difficult in an online environment [11]. During online exams, students have easier access to external resources, such as notes, internet searches, or even help from other people, increasing the likelihood of cheating [27,28].

A new challenge that has emerged during online examinations is the use of artificial intelligence (AI)

systems, such as ChatGPT, by students. These systems can provide quick and accurate answers to test questions, presenting a significant challenge for teachers. The possibility of AI-supported cheating, where students use these tools in real-time to generate responses, raises new security concerns. Educators must address not only traditional forms of cheating but also the misuse of technological tools, which requires continuous improvement in monitoring systems.

Examiners attempt to minimize this risk through various technologies, such as online proctoring software, biometric identification, and AI-driven algorithms that analyze and monitor the behavior of examinees [29,30]. These systems can detect suspicious activities, such as when a student leaves the screen, opens multiple windows, or communicates with others [31, 32]. Despite these technologies, there are still many opportunities for cheating, and not all solutions are 100% reliable.

To prevent cheating, examiners often use different exam formats, such as open-book tests, where students can access resources, but due to the complexity of the questions, the advantage of cheating is minimized [33]. Additionally, tasks requiring individual creativity and problem-solving skills are becoming more popular, as they are less likely to be easily answered using external resources [34].

Verifying the identity of students is also a critical issue. Many institutions use biometric identification, such as facial recognition or voice-based verification, to ensure that the exam is indeed being taken by the authorized student [35,36]. However, there is still the possibility that someone else may take the exam on the student's behalf, especially if there is no constant camera monitoring.

In conclusion, while the spread of online exams is accompanied by numerous technological tools to prevent cheating, these methods are not without challenges [37]. The continuous development of appropriate solutions and exam formats is necessary to ensure the long-term integrity and fairness of online exams.

III. Best Practices and Practical Examples

Numerous educational institutions participate in the administration of online exams, which were introduced due to the COVID-19 pandemic and have been maintained since then for the convenience of examinees. These institutions implement security measures and technological solutions to ensure the integrity of the exams and prevent cheating. Below, we present examples of how online exams are conducted at various language exam centres and universities.

A. Online Language Exams

In online exams conducted by ITOLC [38] (International Test of Language Competence), proprietary exam software is required to be installed on the primary device. Identity verification is mandatory, requiring two forms of identification, and the exam is monitored using a dual-camera system. Students use a desktop computer or laptop to take the exam while a second device (e.g., mobile phone) monitors the examinee's environment to prevent cheating. The exam software includes a chat feature that allows communication with the online proctor at any time.

Euroexam [39] also employs a dual-camera monitoring system and uses biometric identification to ensure the examinee's identity. Identification is based on facial and keystroke patterns, ensuring that the exam is taken by the authorized individual. Proctors monitor the examinee's face and exam environment, as well as their activities on the screen, ensuring the integrity of the exam. The Safe Exam Browser software is used to prevent the examinee from launching other programs or conducting internet searches during the exam.

TELC [40] (The European Language Certificates) online exams also operate with a dual-camera system: one camera monitors the examinee's face, while the other observes the screen and keyboard. This method ensures that students cannot seek external assistance during the exam. TELC places significant emphasis on identity verification, conducted through both biometric and document-based checks, and utilizes the Safe Exam Browser.

The LanguageCert [41] exam center has also introduced online exam options, allowing examinees to take exams from the comfort of their own devices at home. The exam is monitored continuously by cameras and other security software (Safe Exam Browser) to ensure exam integrity. Students will also need to use the ZOOM application. LanguageCert exams require prior verification of technical requirements and the examinee's identity to guarantee a fraud-free exam environment.

The Origo Language Exam Center [42] provides online exam options as well, implementing similar security measures. Examinees log in from their own devices, and during the exams, they are under constant supervision. Students participate using a camera and microphone, monitored by Origo's proctors. The Safe Exam Browser software is responsible for security and fraud detection. The examinee must complete the identification process 72 hours before the exam.

The Corvinus Language Exam Center [43] also offers online exams, during which examinees ensure that no other programs can run using the Safe Exam Browser. Students undergo a technical check to verify their exam environment and devices in advance. Corvinus also uses a dual-camera monitoring system to observe the examinee's activities and surroundings.

The IELTS [44] (International English Language Testing System) also offers online testing options, using the Inspera Exam Portal software. This test can be taken from home and is monitored continuously by cameras to ensure the authenticity of the exam. In the case of IELTS, strict identity verification is performed alongside the fulfillment of technical requirements to prevent misuse. Human and AI-based exam supervision is implemented.

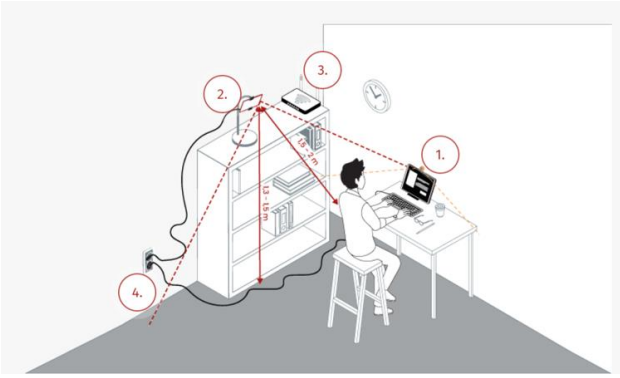


Figure 1. Setup of Online Exam Location *(source: https://telc.hu/nyelvvizsga/otthoni-online-vizsga/)*

The administration of online exams across different language exam centers follows similar principles: dual-camera supervision (figure 1), biometric or document-based identification, and the use of security software to prevent cheating. The need for online testing arose due to the COVID-19 pandemic, but based on experience, this option has remained available in many places, providing greater flexibility for students.

B. Online Exams at Universities

Before the pandemic, most universities predominantly used traditional paper-based exams and in-person evaluation methods. However, during COVID-19, educational institutions were forced to implement online systems to ensure the safety of students and faculty [45]. This sudden transition allowed students to take exams from their homes, using various devices, making the process more flexible and, in many cases, more efficient. Universities issued specific guidelines for conducting online exams, particularly outlining the procedures for both oral and written exams [46].

After the pandemic subsided, several universities decided to retain the option of online exams [47]. This option was primarily maintained for oral exams, which were conducted via video conferencing tools (such as Zoom, Skype, MS Teams) along with identification verification methods for students. For written exams, the continuation of online testing was less common, but when retained, universities supplemented their existing e-learning platforms (Moodle, Canvas, CooSpace) with proctoring software such as Safe Exam Browser.

In 2021, the Faculty of Economics at Eötvös Loránd University (ELTE) established a modern exam center through a 350 million HUF investment. This facility is considered groundbreaking in Hungarian higher education, as it enables the administration of 10,000 exams per week with state-of-the-art technology and controlled conditions. The center can accommodate 200 students at once, providing monitors, keyboards, and drawing tablets for use during exams. The monitors are equipped with privacy screens to prevent visibility from adjacent screens, and all exam activities are recorded to prevent cheating and enhance the effectiveness of assessments. The center is capable of administering up to 200,000 exams per year, contributing to a significant reduction in paper-based testing, saving over one million sheets of paper annually. This investment aims to make ELTE's economics program even more attractive to students by offering innovative, environmentally friendly solutions for assessments [48].

C. Students' Opinion on Online Exams

tudents' opinions on online exams were generally positive, particularly in terms of flexibility and convenience. Research conducted among medical students, for example, showed that while they were satisfied with the format of online education and exams, many missed the personal interaction present in traditional education [49]. Some students, therefore, preferred a hybrid educational model, where online and in-person elements are combined. At the same time, online exams allowed them to participate from the safety and comfort of their home environment, making the exam period smoother for many.

A study comparing the performance of students in economics courses at Óbuda University found that students performed better in online exams compared to traditional paper-based tests [50]. In the same study, there was no gender difference in the results of online exams, but part-time students outperformed full-time students. Other research indicated that the average scores for online exams were slightly lower than traditional exams, but the difference was not statistically significant [51]. It is important to note that a significant portion of students gave positive feedback on the administration of online exams. During the COVID-19 exam periods and afterward, students appreciated the flexibility offered by online exams, especially in terms of scheduling and location independence.

According to the results of the COVID-19 Global Student Survey, Hungarian students ranked among the top in the world in reporting that they were able to improve their performance in the new online learning environment [52]. Overall student satisfaction was also exceptionally high regarding the administration of online exams and the presentation of course materials by instructors. Most students found the time allocated for online exams to be adequate and were satisfied with the difficulty level of the tests.

In summary, research shows that students' performance in online exams is comparable to traditional exams, and in many cases, they were satisfied with the flexibility and opportunities provided by the new method, although some missed the personal interaction.

IV. CONLUSION

The emergence of online exams is closely linked to the digitalization of education, which began in the early 21st century. Technological advancements and the spread of e-learning platforms have enabled educational systems to continually adapt to the new digital environment. However, online exams truly gained momentum due to the COVID-19 pandemic when traditional educational institutions were forced to shift both teaching and testing to remote platforms. During the pandemic, the demand for online education surged dramatically, along with the need for online exams. Various technological solutions, such as Zoom, Google Meet, and e-learning platforms (Moodle, Canvas), became essential tools in the online teaching and examination processes. Online exams provided students the

opportunity to participate in exams from the safety of their homes without interrupting their studies.

One of the main advantages of online exams is flexibility, allowing students to take exams at a time and place that suits them best. However, several challenges have arisen, such as preventing cheating, managing technical issues, and verifying students' identities. Educational institutions have implemented various technological solutions to address these challenges, including dual-camera systems, online proctoring software, and AI-driven algorithms. While the COVID-19 pandemic played a key role in the rapid spread of online exams, many institutions have continued to offer this format as it provides greater flexibility and accessibility for students. The benefits of online exams, such as cost-effectiveness and an inclusive learning environment, continue to play an important role in the future of education.

One of the most critical aspects of implementing online exams is maintaining integrity, preventing cheating, ensuring equal access, and securely identifying students. To this end, biometric identification is recommended, allowing for face and voice verification both before and during the exam. This solution helps to authenticate the student's identity, minimizing the chances of someone else taking the exam on their behalf. For technical security, it is advisable to mandate the use of proctoring software, which locks the test-taker’s computer, preventing the use of external programs or searches. Additionally, a dual-camera system that monitors the test-taker and their surroundings can effectively prevent cheating, making it particularly useful when students take exams from home.

During online exams, open-book formats are also suggested, as they minimize the benefits of cheating by emphasizing problem-solving skills and creativity. Randomized question sets can further prevent the sharing of answers among students, ensuring the exam's integrity. Before exams, it is crucial for students to familiarize themselves with the exam system through technical trial sessions, and institutions should provide continuous technical support during the exam. Equal access must also be ensured, for instance, by offering equipment loans or allowing students to use university exam centers if they lack adequate technological resources.

ELTE’s example of establishing a modern exam center demonstrates the importance of institutions creating facilities that provide the necessary technical conditions for online exams. Such solutions can make online exams safer, fairer, and sustainable in the long term.

In our study we highlighted the administrative benefits of online exams, emphasizing flexibility, accessibility, and inclusivity, particularly for non-traditional students. The significant technological challenges, including the need for stable internet access, effective identity verification, and sustainable support systems to ensure fair participation are crucial. Finally, our research underlined critical ethical considerations, such as safeguarding academic integrity, privacy, and balancing these aspects with AI tools, all crucial for maintaining trust in digital examinations.

V. REFERENCES

[1] T. Kőfalvi, E-tanítás. Információs és kommunikációs technológiák felhasználása az iskolában. Budapest: Nemzeti Tankönyvkiadó, 2006.

[2] A. Benedek (szerk.), Digitális pedagógia 2.0. 2. kiadás, Budapest: Typotext, 2013.

[3] A. Benedek, “Távoktatás másként!!! – A digitális kor pedagógiai kihívásaihoz,” Opus et Educatiom 7. évf., 3. sz., 2020.

[4] K. Szabó, T. Juhász, M. Kenderfi, “Felsőoktatás a covid-19 árnyékában – hazai tapasztalatok oktatói oldalról,” Vezetéstudomány / Budapest Management Review 3 LIII. ÉVF. 2022. 6. SZÁM / ISSN 0133-0179 DOI: 10.14267/VEZTUD.2022.06.01

[5] J. Ollé, “Egy módszer alkonya: katedrapedagógia végnapjai a felsőoktatásban,” In: Dobó-Perjés-Temesi (szerk.) Korszerű felsőoktatásipedagógiai módszerek, törekvések. Konferencia előadások. NFKK Füzetek5. AULA Kiadó Kft., BCE. pp. 22-31. 2010. ISBN 978-966-89082-09http://unipub.lib.uni-corvinus.hu/1231/1/NFKK_5_vegleges.pd

[6] J. Vereckei, "Személyes jelenlét nélküli felsőoktatás a vizsgaszervezés aspektusából a közoktatási vezetőképzésben," Opus et Educatio, 8.2., 2021.

[7] I. Szenciová, K. Kántor, S. Darvay, E. Tothova Tarova, M. Nagy, Z. Fehér, P. Balázs, “Az online oktatás bevezetése és hatékonysága a Selye János Egyetemen a covid-19 pandémia első hulláma után,”. Conference: 13th International Conference of J. Selye University, December 2020.

[8] I. Miskolczi, “Elenőrzési és vizsgáztatási lehetőségek bővítése a MOODLE rendszerben,” Hadmérnök, V. Évfolyam 3. szám, 2010. szeptember

[9] A. Dávidovics, “Online oktatás és vizsgák a világjárvány alatt a pécsi orvostanhallgatók szemszögéből,” Porta Lingua. 2023/1. pp. 107-118. 2023. DOI: https://doi.org/10.48040/PL.2023.1.10

[10] M. Amirhosein, B. Beszedes, I. Felde, L. Nadai, G. Nima, “Electrical characterization of CIGS thin-film solar cells by two- and four-wire probe technique,” Modern physics letters B 2020, 16 p. 2020.

[11] M. Castro, S. Manoharan, U. Speidel, X. Ye, J. Zu, ”Observations of Cheating Behaviours in Online Examinations and Tools for Mitigation,” 2023 IEEE Frontiers in Education Conference (FIE). 2023.

[12] T. Fatima, F. Azam, A. W. Muzaffar, “A Systematic Review on Fully Automated Online Exam Proctoring Approaches,” 2022 24th International Multitopic Conference (INMIC), 2022.

[13] D.P.D. Dayananda, S. Vasanthapriyan, K.G.L. Chathumini, S. Fernando, “A Systematic Literature Review on Online Exams in COVID-19 Pandemic: Assessment Methods, Students' Preferences, Dishonest Behaviors and Challenges in Online Exams,” 2021 From Innovation To Impact (FITI), 2021. doi: 10.1109/FITI54902.2021.9833059

[14] N. Tolner, M. Pogátsnik és J. Módné Takács, “A mesterséges intelligencia szerepe az online vizsgáztatásban,” Iskolakultúra, 33(10), pp 39–55. 2023. https://doi.org/10.14232/iskkult.2023.10.39

[15] D. Chuanmei, C. Simin, L. Hui, “Young children’s online learning during COVID-19 pandemic: Chinese parents’ beliefs and attitudes,” Children and youth services review, 118: 105440, 2020.

[16] Cs. Kézi, “Globalization in Higher education, Advantage and Disadvantage of Online Education,” International Journal of Engineering and Management Sciences, 8(4), pp. 1-9, December 2023. DOI: 10.21791/IJEMS.2023.028

[17] A. Lomellini, P. R. Lowenthal, C. Snelson, and J. H. Trespalacios, “Higher education leaders’ perspectives of accessible and inclusive online learning,” Distance Education, 43(4), pp. 574–595., 2022. https://doi.org/10.1080/01587919.2022.2141608

[18] S. Burgstahler, “Creating Online Instruction that is Accessible, Usable, and Inclusive,” The Northwest eLearning Journal, 1(1). May 2021.

[19] R.T. Thorat, N. Poornima, R. S. Kamat, “Revolutionizing Online Exams on a Budget: A Zero-Cost, Open- Source Approach for Challenging Times,” International Conference on India's Role in the Global Scenario (Education, Technology, Business and Climate) At: CSIBER, Kolhapur, October 2023.

[20] S. N. M. Zaihirain and V. Selvarajah, "A Technical Analysis on the Current Implementation of Online Examination System and the Factors Influencing the Issue," 2022 IEEE 2nd International Conference on Mobile Networks and Wireless Communications (ICMNWC), Tumkur, Karnataka, India, pp. 1-5, 2022. doi: 10.1109/ICMNWC56175.2022.10031940.

[21] L. T. Parázsó, Online teszt és tudásszintmérés, In: Agria Media 2011: Az információs társadalom az alkotó tudás társadalmának digitális előtere, Eszterházy Károly Főiskola Líceum Kiadó, pp. 117-125, 2011.

[22] M. Kuleva and O. Miladinov, "Exploring the efficacy of online proctoring in online examinations: a comprehensive review", ETR, vol. 2, pp. 192–196, Jun. 2024. doi: 10.17770/etr2024vol2.8058.

[23] P. J. Sam, O. Farhath, and M. P. Selvan. "AI based Proctoring System–A Review," 2024 International Conference on Inventive Computation Technologies (ICICT). IEEE, 2024.

[24] B. Mistry, H. Parekh, K. Desai, N. Shah, "Online Examination System with Measures for Prevention of Cheating along with Rapid Assessment and Automatic Grading," 5th IEEE International Conference on Advances in Science and Technology, ICAST 2022, pp. 28–34, 2022.

[25] M. Nicolau, "BSGenerator: a Tool to Design Unique Online Examinations using Question Variations.," 2022 12th International Congress on Advanced Applied Informatics (IIAI-AAI), pp. 182–187, 2022.

[26] F. Skopljanac-Macina, I. Zakarija, B. Blaskovic, "Organizing Online Exams during the COVID-19 Pandemic," 2021 44th International Convention on Information, Communication and Electronic Technology, MIPRO 2021 - Proceedings, pp. 783–788, 2021.

[27] D. R. E Cotton, P. A. Cotton, és J. R. Shipway, "Chatting and cheating: Ensuring academic integrity in the era of ChatGPT," Innovations in Education and Teaching International, 61. 2 sz., pp. 228–239. 2023.03.13. Online: https://www.tandfonline.com/doi/full/10.1080/14703297.2023.2190148#d1e293 (download: 2024.05.29.)

[28] M. Srikanth, R. Asmatulu, "Modern cheating techniques, their adverse effects on engineering education and preventions," International Journal of Mechanical Engineering Education. 42. 2 sz., pp. 129-140. 2014. Online: https://www.researchgate.net/publication/276387081_Modern_Cheating_Techniques_Their_Adverse_Effects_on_Engineering_Education_and_Preventions (download: 2024.05.29.)

[29] S. P. Saurav, P. Pandey, S. K. Sharma, B. Pandey, R. Kumar, "AI Based Proctoring," Proceedings - 2021 3rd International Conference on Advances in Computing, Communication Control and Networking, ICAC3N 2021, pp. 610–613, 2021.

[30] A. Shata, Z. N. Haitaamar, A. Benkrid, "Baseline Estimation in Face Detection for AI Proctored Examinations through Convoluted Neural Networks," 2023 International Conference on IT Innovation and Knowledge Discovery, ITIKD 2023.

[31] P. D. Gordon and M. P. McDougall, "A Tool for the Discovery of Academic Misconduct in Online Assessments Using Student Activity Logs," 2023 ASEE Annual Conference & Exposition, Baltimore , Maryland, ASEE Conferences, 2023.

[32] A. Gupta, A. Bhat, "Bluetooth Camera based Online Examination System with Deep Learning," Proceedings - 2022 6th International Conference on Intelligent Computing and Control Systems, ICICCS 2022, pp. 1477–1480. 2022.

[33] K. Quille, K. Nolan, B.A. Becker, S. McHugh, "Developing an Open-Book Online Exam for Final Year Students," Annual Conference on Innovation and Technology in Computer Science Education, ITiCSE, pp. 338–344, 2021.

[34] I. Pšenáková és M. Štrbo: Online is lehet „puskázni" = Online Can also Be „Cheating". Mesterséges Intelligencia, 4 (1). pp. 49-55. 2022.

[35] T. Nguyen, L.N. Vo Phi, "LA4OX - A Live Authentication for Online Exam Quoc," Proceedings - International Conference on Knowledge and Systems Engineering, KSE 2023.

[36] H. Wang, X. Yan, H. Li, F. Xu: Double-frame rate-based cheat detection. Proceedings of SPIE - The International Society for Optical Engineering, 2023

[37] Gy. Györök, B. Beszédes, "Artificial Education Process Environment for Embedded Systems," In: Orosz, Gábor Tamás (szerk.) 9th International Symposium on Applied Informatics and Related Areas - AIS2014 Székesfehérvár, Magyarország, Óbudai Egyetem, pp. 37-42, 2014.

[38] ITOLC Language Exam Center. https://itolc.hu/online-nyelvvizsgazas/ (download 2024.07.18.)

[39] Euroexam Language Exam Center. https://www.euroexam.org/hu/vizsgainfok/hungary/otthoni-online-nyelvvizsga/ (download 2024.07.18.)

[40] TELC Language Exam Center. https://telc.hu/nyelvvizsga/otthoni-online-vizsga/ (download 2024.07.18.)

[41] LanguageCert Language Exam Center. https://languagecert.hu/online/ (download 2024.07.18.)

[42] Origo Language Exam Center. https://www.onyc.hu/origo-online-vizsga/ (download 2024.07.18.)

[43] Corvinus Language Exam Center. https://bce.nyelvvizsgajelentkezes.hu/doc/BCE_T%C3%A1j%C3%A9koztat%C3%B3_vizsg%C3%A1z%C3%B3nak_az_online_nyelvvizsg%C3%A1kr%C3%B3l.pdf (download 2024.07.18.)

[44] IELTS. https://www.britishcouncil.hu/angol-nyelvvizsga/ielts/datumok-dijak-helyszinek/ielts-online/ (download 2024.07.18.)

[45] F. Škopljanac-Mačina, I. Zakarija, and B. Blašković. "Organizing online exams during the COVID-19 pandemic." 2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO). IEEE, 2021.

[46] E. O. Bereczki, L. Horváth, O. Kálmán, K. Káplár-Kodácsy, H. Misley, A. Rausch, Z. Rónay, and K. Solymosi, "Kiegészítő segédanyag a hallgatói munka értékeléséhez a távolléti oktatásban megvalósuló kollokviumok és vizsgák esetében," Budapest, ELTE-PPK. 2020.

[47] L. Furcsa, "Az online szóbeli vizsgáztatás észlelt előnyei az egyetemeken," In: E. Berghauer-Olasz, É. Csopák, I. Greba és K. Lizák (szerk.): Krízishelyzetek hatása és kihívásai az oktatásban, Nemzetközi tudományos konferencia Beregszász, 2023. március 30–31.

[48] https://infostart.hu/belfold/2021/11/04/uj-korszakot-indithat-a-felsooktatasban-az-elte-vizsgakozpontja-kepek (download 2024.05.30.)

[49] A. Dávidovics, "Online oktatás és vizsgák a világjárvány alatt a pécsi orvostanhallgatók szemszögéből," PORTA LINGUA (1). pp. 107-118, 2023.

[50] Zs. Deák, "Óbudai Egyetem hallgatóinak gazdaságtan kurzusokon elért teljesítményének longitudinális vizsgálata papír alapú versus online alapú tesztekben," In: Á. Csiszárik-Kocsir, J. Varga, M. Garai-Fodor (szerk.) Vállalkozásfejlesztés a XXI. században 2022/2 kötet - Globális reakciók, lokális akciók a gazdaság rezilienciájának erősítéséhe,. Budapest, Magyarország, Óbudai Egyetem, Keleti Károly Gazdasági Kar, 367 p. pp. 103-112, 2022.

[51] T. K. Baranyai, E. Debrenthi, "Az online matematikatanítás tapasztalatai," PedActa 10. kötet, 2. szám, pp. 33-42. 2020. https://doi.org/10.24193/PedActa.10.2.4

[52] CovidSoclab.org: Global Student Survey http://www.covidsoclab.org (download: 2024.07.25.)

Where are we with Micromobility?

Richárd Veres
Alba Regia Faculty, Obuda University
Doctoral School of Innovation Management OU
Szekesfehervar, Hungary
veres.richard@amk.uni-obuda.hu

Abstract - In this short paper, I wanted to look at the megatrends and trends in micromobility, and how forecasts are evolving. What strategies, programmes and initiatives have emerged in the European Union to support the introduction of more sustainable mobility solutions, and what good examples can be highlighted. I summarised the experience of Hungarian cities with micromobility. What sharing-based e-scooter providers have spread their services to larger municipalities, where is the current state of play in the deployment of Hungarian public bicycle systems. A SWOT analysis was used to collect possible outcomes of the introduction of urban solutions for micromobility and to outline scenarios.

Keywords — micromobility, trends, innovation, sustainability, urban transport.

INTRODUCTION

As cities develop and become more populated, new technologies emerge in various fields, rewriting the rules of the game in their innovative essence, disrupting people from their usual rhythm and generating alternative solutions to the challenges of everyday life. Smart urban development and management is becoming increasingly important with urbanisation, which seeks to increase economic efficiency, improve the quality of local public services, protect the built and natural environment, and use state-of-the-art information technology tools to implement developments with the involvement of residents in a sustainable way (Lechner Knowledge Centre) [1]. City leaders should continuously explore the integration of proven innovations to address the challenges of congestion, air and noise pollution, climate change impacts, road safety and parking. There is a growing demand from city dwellers for a liveable city, which goes hand in hand with increasing green spaces for recreational activities, solving the transport problems that currently affect almost all large cities, providing easier access to public services, and a vision of a city that already prioritises environmental awareness and sustainability.

In this study, I narrow down the broad range of people's needs to the issue of micromobility, which is present in the transport of all large cities and generates social and economic challenges. Anyone living in a large city who usually encounters some form of micromobility (most often bicycles, e-scooters, etc.) will usually have an opinion on whether or not there is room for it in local transport management. Many like it, many would banish it, but before looking at its situation in Hungary, it is necessary to look at what trends are influencing the organisation and development of urban life in this direction.

MEGATRENDS AND TRENDS

What megatrends are in place that can shape the cities of the future in the long term, which we want to make greener, more sustainable and more liveable for all. Based on the European Commission's megatrends page (European Commission, The Megatrends Hub, 2024), I have identified the driving forces that are relevant to my topic [2].

Continuous urbanisation, one of the main drivers of urban population growth - expected to reach 5 billion people by 2050 - is pushing decision-makers to look for smart city solutions to make cities more sustainable, liveable and intelligent. Within the megatrend, there is also a trend towards the interplay between *technology and the city*, as digitalisation enters different levels of operation, impacting on health, mobility, citizen participation in decision-making, community building. There is also a *trend towards the rise of* various *green actions*, with local city leaders and civil society working together to do something for their environment, to reduce their own carbon footprint and to achieve the SDGs. There is also an emerging trend for *cities to organise themselves into partnership networks* , with more and more municipalities sharing their experience in urban management and development, passing on the knowledge they have acquired and thus speeding up the resolution of various problems.

Another megatrend in urban life *is the acceleration of technological change and hyper-connectivity*, which also means that the range of smart city solutions available to impact on citizens' daily lives is constantly expanding. People are constantly 'generating' data through *their wearable or other smart devices* about themselves, their habits, their mobility, their health, access to which *gives data controllers great power and also poses great risks.* This data collected has the potential to create personalised solutions to improve quality of life, but on the downside, it can also be the basis for cyber-attacks. The transformation of travel habits and the *mobility of the future is a trend (and the twin transition)* that is manifesting itself today in the proliferation of micromobility devices and the spread of autopilot systems.

Urban growth *also increases consumption*, which can also be identified as a megatrend. Nowadays, manufacturers and service providers are trying to reach people at all levels of society in order to make more profit through various online marketing campaigns and online

shops. *The* need for *sustainable consumption* is already emerging as a trend among a group of consumers, which is influencing their purchasing decisions and influencing decision-makers in business organisations to take sustainability into account in the global competition for consumers.

GARTNER'S FORSIGHT

Each year, Gartner presents the latest technologies and digital innovations on a hype cycle, which, when aggregated by different areas, indicate their maturity and acceptance. This summary chart is also published every year for the "Smart City" theme, to help clients (urban developers, entrepreneurs and decision-makers) to identify the technologies that will be the most appropriate for their developments. The Gartner Hype Cycle helps them understand the risks and opportunities of innovation by categorising technologies, digitalisation drivers and strategies into one of five stages of development. The Smart City Technologies and Solutions cycle (2023) includes the mobility-related Last Miles Delivery Solutions, which are increasingly recognised as important. Decision-makers are looking at how to integrate last-mile delivery micromobility devices into the local transport system. The early users of the technology are sharing their experiences to help them understand this process. The technology's place in the diagram indicates that it could be integrated and enhanced in day-to-day transport management within a year or two, both along individual and urban strategies. "Shared Mobility" is in the same phase, with more and more users understanding the benefits of the application and seeing the potential to integrate it into their daily routine.

Gartner's researchers indicate that city managers can look for these solutions as part of their sustainable urban development strategy to steer development towards a smarter city, mitigate the negative impacts of climate change and help their citizens to achieve more environmentally conscious, sustainable mobility [3].

ABOUT MICROMOBILITY

According to a Deloitte survey, both city residents and city authorities are looking for a change in urban transport to reduce congestion and carbon emissions and move towards more sustainable mobility modes (2deloitte.com, 2023) [4]. Micromobility is being explored as a possible innovative solution to the problems caused by continued urbanisation and the growing mobility needs and demands of society. Why this might be a good answer is explored by Horace Dediu, founder of the global media platform Micromobility Industries, who is credited with coining the word "micromobility" itself back in 2017 (Horace Dediu, 2019) [5]. With the number of micromobility devices outnumbering electric cars, growing rapidly, cost-effective transportation, reducing emissions, contributing to sustainability goals, the micromobility interface is poised for further development (Horace Dediu, 2023) [6].

What can be considered micromobility? According to Berényi et al., due to the continuous change of the means of transport, we should aim at creating a flexible framework, where common characteristics are urban transport means that allow micro-trips, travel at low speed, are small in size and weight, can be human-powered or electrically driven, owned or rented (Berényi et al., 2022) [7].

According to McKinsey researchers, shared mobility is gaining ground in urban transport, driven by three forces: first, the shift from individual vehicle users to car sharing; second, the shift from individual to autonomous driving; and third, the shift from larger to smaller vehicles (McKinsey. com, 2023) [8]. Based on these findings, four segments of shared mobility (hailed mobility, shared micromobility, car sharing, urban aerial mobility) have been highlighted, which could generate up to USD 1,000 billion in consumer spending by 2030. Recommendations to car manufacturers include considering entering the shared mobility market as car ownership declines. A good example of a move in this direction is Lancia, the Italian carmaker that is entering the electric bicycle market and will focus on micromobility in the future (Trendhunter.com, 2022) [9]. Many car brands will follow suit. Sales of electric bicycles in North America and Europe will almost double after Covid by 2022 (2deloitte.com, 2022.) [10]. In a similarly interesting shift in consumer habits within the micromobility sector, the company Segway decided to replace its leading product "segway" with a scooter in 2017 (Segway. com, 2017) [11]. The reasons may include the fact that the segway has been popularised as a means of enhancing the experience of exploring tourist destinations (Michalkó, 2023) [12] and has become one of the most modern means of transport for sightseeing (Miskolczi, 2023), but has failed to fulfil the role it was intended to play in general urban mobility [13].

The development and integration of electric drivetrains and micromobility solutions are fundamentally tied to advancements in Embedded Systems, power management systems, and Electro-Pneumatic Systems. These fields are critical for optimizing the performance, efficiency, and safety of electric mobility technologies. In the context of higher education, these topics offer a rich, interdisciplinary foundation that can be effectively utilized as comprehensive teaching material. Incorporating these subjects into the curriculum allows students to engage with real-world, hands-on applications, preparing them for the challenges of emerging technologies in sustainable transportation and enabling a deeper understanding of the technical and engineering principles driving future innovations in electric mobility [14].

Shared on-demand mobility in megacity travel patterns is expected to grow from 8% to 23% between 2021 and 2035, according to a forecast in Statista's statistical report on the market environment for micromobility start-

ups (Statista.com, 2024). The micromobility unicorns continue to grow in value, adding another round of investment as they evolve, with Bolt ($8.4bn), G7 ($2bn), Tier Mobility ($2bn), Rad Power Bikes ($1.4bn), Segway-Ninebot and Voi Scooters ($1bn) in the 2023 order. Tier Mobility, which is also present in our country, had 2020 revenues of €57.3m [15]. The micromobility market is still growing, but there are likely to be disruptions, as illustrated by the protests abroad in Paris, Los Angeles and Dallas (bans, demonstrations, etc.), which could lead to a more predictable pattern of activity and operations for companies providing innovative transport solutions.

EU INITIATIVES, PILOTS, GOOD PRACTICES

The driving forces behind the growth of the micromobility sector are the travelling public itself, as well as nation states committed to addressing climate change challenges and contributing to the achievement of the UN Sustainable Development Goals (SDGs). The European Union supports and funds these goals through flagship projects. A good example is the launch of the "Scalable Cities" programme, which aims to create innovative, sustainable urban communities in smart and climate-neutral European cities, working with cities, academia, industry, associations and consultants (European Commission, Scalable Cities, 2022) to create shareable good practices through pilots that can be implemented in large cities [16].

In its "Sustainable and Smart Mobility Strategy", the European Commission has set out 82 initiatives for 4 years, one of the 10 "flagships" of which is the healthy and sustainable mobility of people between cities and towns, with micromobility (e.g. improving the infrastructure conditions for cycling) as one of the elements to create a connected, seamless and sustainable transport mix [17]. One of the defining concepts of European urban mobility policy is the Sustainable Urban Mobility Plans (SUMPs), presented in 2013 and updated in 2023 in line with new EU strategies and policy priorities . The European Commission is calling for the adoption of the SUMP concept in European cities to improve the overall quality of life of city dwellers, providing a framework for integrating innovation and new mobility services (European Commission, Sustainable Urban Mobility Plans (SUMPs), 2023) [18].

The aim of the European Institute of Innovation and Technology (EIT) Urban Mobility programme is to encourage people to change their transport habits through a range of innovative project initiatives to create more liveable urban spaces [19]. Through an integrated approach involving all urban mobility actors, industry and academic partners, and by reaching out to city dwellers, they aim to turn cities into laboratories to solve real problems with new technologies. Examples include the MiGriS ("Micromobility Grid Solutions") or the Parkedbyme projects (EITUrbanMobility.eu, 2024).

The European Commission is supporting the implementation of the European Green Deal and the shift to sustainable urban mobility through the CIVITAS flagship initiative. As a "Network of Cities", it helps cities to organise their professionals in a knowledge-sharing network, to facilitate the learning process from the experience of pioneering cities in finding solutions, to provide opportunities for development, testing and preparation of innovative sustainable urban development actions (e.g. R&D projects in the field of smart and sustainable urban transport, exchanges of professionals) [20]. The Hungarian CIVINET was established in 2016 and has recruited a number of municipalities who are working together with professional members (NGOs, research institutes, universities) to find forward-looking urban transport solutions. Within CIVITAS, the SPROUT project (Sustainable Policy RespOnse to Urban mobility Transition) should be highlighted, which aims to provide a sustainable policy response to the mobility transition in response to the rapidly changing urban transport environment [21]. They also operate an open innovation community platform through which many more stakeholders' views can be expressed. Budapest is among the cities participating in the project.

Of course, there are initiatives not only from the European Union, but the mobility industry itself would like to see this explosive growth, which has taken place in just a few years, become sustainable, while achieving the expected benefits (reduced carbon emissions, congestion, reduced air pollution, more liveable urban areas), by supporting the development of a regulatory framework and encouraging the use of technologies. The "Micro-mobility for Europe" coalition was created by micro-mobility service providers (Bird, Bolt, Tier, Lime, etc.) to create a more sustainable transport ecosystem [22]. To this end, they work together (e.g. by providing data) with decision-makers at local, national and European level to regulate, plan and develop infrastructure more effectively towards the common goal of more liveable cities (e.g. support for targeted cycling infrastructure, creation of a micromobility day, development and harmonisation of legal frameworks, multimodal ticketing) (Micromobilityforeurope.eu, 2024).

The spread of youthful, trendy mobility "festivals" helps to raise awareness of new technologies, share experiences of technologies, showcase good practices, educate the public and build a community of support for alternative mobility tools. One such festival is the atypical international shared mobility conference "Shared Mobility Rocks" (Shared-mobility.rocks, 2024) [23]. In addition, one of the biggest micromobility conferences is "Micromobility Europe", organised in Amsterdam in 2024 with a parade of the industry's key founders, brands, manufacturing companies, metropolitan experts, start-ups, operators, developers, as well as key trends, technological innovations and product launches (Micromobility.io, 2024) [24].

MICROMOBILITY EXPERIENCES IN HUNGARIAN CITIES: E-SCOOTERS

In Hungarian cities too, decision-makers are exploring how to respond to the needs of local society by introducing digitalisation and smart solutions to improve the quality of life, in line with the implementation of sustainable urban development strategies. One of these areas of innovative, smart and accessible developments is micromobility, which carries the vision of a greener, pollution-free, healthier city of the future. With renewed, more sustainable public transport organised in intermodal hubs, free from congestion, complementing last-mile mobility, channelling services from sharing economy actors, it can create a framework that can free cities from excessive infrastructure demands (required by the mass of cars parked in public spaces) and give people back their walkable and community-building spaces. A major advantage of the vehicles available from the sharing economy providers is that the facility can be used by more people in a day, as opposed to the daily under-utilisation of private cars.

In Hungarian cities, cyclists have been joined by micro-mobility users (e-scooter, solowheel, e-bike), and cities have decided to make mobility products available to sharing economy operators. In Hungary, for the first time in 2019, Lime Technology Ltd. offered the first platform-based e-roller sharing service in Budapest. The Covid epidemic has boosted digitalisation and automation and, in turn, has influenced people's transport habits as they seek greater safety in their mobility. This gave a boost to the spread of the service in Hungary, with the leaders of major cities taking the decision to open their doors to Lime Technology Kft. Székesfehérvár (2021), Szeged (2021), Veszprém (2021), Szászhalombatta (2021), Szombathely, Szolnok (2022), Érd (2022), Nyíregyháza (2022), Hódmezővásárhely (2023), Sopron (2023), Agárd (2021), Gárdony (2021), Velence (2021) and the towns around Lake Balaton (2021). In the summer of 2021, Bird Hungary entered the Budapest ride-sharing market, followed by TIER Operations Hungary Kft in September 2021 and Whoosh e-roller in 2023. Bird Hungary continued its expansion in 2023 in Békéscsaba, Dunaújváros, Eger, Nagykanizsa and Zalaegerszeg. In 2023, TIER launched in Győr, Miskolc and Balaton. In 2023, HOPP Mecsek Kft. started operations in Pécs, with Icelandic operator HOPP. Eight of the 25 cities with county rights do not currently have e-roller services. The advent of e-roller technology in people's daily mobility practices is both a challenge and a relief. On the one hand, the possibility of flexible and free parking is a very practical feature and an advantage for users because of the ease of access to transport hubs and the ability to cover the last mile. On the other hand, for pedestrians, this feature tended to be a nuisance, an annoyance and a disadvantage. This problem could be solved by regulation, as has been done in Budapest by designating Mobi Points (Micromobility Points) where devices can be dropped off (on-street storage). Operators are demanding regulation and are willing to work with local authorities to develop a framework (operation, management, parking zones) that meets local needs. They also make recommendations for the development of regulation, as Lime has done in its 10 points for policy makers (Li.me/blog, 2023), to create a proper micromobility programme for business development and growth environment based on security [25]. Szemere et al. investigated the opinion of the user side in Hungary, concluding that users are also waiting for legislation to be developed (speed limit issues, infrastructure use and parking space allocation, etc.) (Szemere et al., 2024) [26].

According to figures from the Future Mobility Association, in April 2023 the number of scooters used in daily transport was estimated to be between 80,000 and 100,000 (about 25,000 rented) (Jovomobilitasa.hu, 2023) [27]. In 2023, the number of devices placed by scooter operators entering large cities could be around 3,000-3,500, given the trial periods. After the designated period, the number could increase or decrease depending on the popularity. There is also an example in Hungary of a railway company (GYSEV Zrt.) operating the scooters (Sopron). It should be noted that electric vehicles are also used to collect and deliver the scooters, thus reducing emissions.

COMMUNITY BICYCLE SCHEMES

Improving cycling infrastructure is not only good for cyclists, but also for the whole micromobility system. The more cycle lanes are built and designated, the more they can be integrated into urban transport without endangering pedestrians.

In Hungary, the deployment and uptake of community cycling schemes is varied. Eleven of the 25 cities with county status have a live service and of course Budapest, where in 2014 the “MOL Bubi” public bicycle service (maintained by BKK) was among the first. Also in 2014, the “HeBi” community bicycle system was introduced in Hévíz, promoting access to the city's tourist attractions. It took five years to launch the Pécs “PécsiKe” Community Bike Scheme in Pécs and the “UniBike” University Bike Scheme in Debrecen in 2019. The country's first public bicycle system was launched in Esztergom in September 2013 under the name “EBI” (Esztergom Bicikli, in cooperation with Neuzer), and was further developed in 2020 with the establishment of the “Mária Valéria Bike”, a modern system linking Hungary and Slovakia, in cooperation with the neighbouring municipality of Párkány. In 2022, Szeged expanded its cooperation with Lime Technology Kft., which now also enabled the sharing of LIME electric bikes in the city. In 2023, Veszprém (“V-Bike”), Kaposvár, Tatabánya (“T-Bike”) and Eger expanded their mobility ecosystem with a community bike system. In two large cities, the number of users did not make use of the option provided to ensure sustainability (lack of demand: the majority of locals use their own bikes), so in Nagykanizsa and Szombathely this shared transport alternative was discontinued. In

Székesfehérvár, an extended agreement between the municipality and Lime Technology Ltd. resulted in the launch of the LIME electric bike sharing programme in 2024. And in Békéscsaba, preparations for the introduction of the "CsaBi" community bicycle system started in 2023.

A SWOT ANALYSIS

I examined the possible outcomes of the introduction of urban solutions for micromobility through a SWOT analysis. What are the infrastructural conditions for integrating these devices into the mobility life of a large city? How can the attitudes of the urban population be shaped, and later changed, towards more sustainable means of transport? How can local decision-makers be encouraged to take unpopular decisions in favour of micromobility, while limiting the use of private cars? These are the questions along which I have explored the strengths, weaknesses, opportunities and threats.

Strengths:

- Reduced pollution, cleaner, healthier living conditions.
- Reduction of congestion on road networks, increase in areas that can be converted into green spaces.
- Reduction in time spent on transport.
- Increase in walkable, community and micromobility spaces.
- Encouraging a health-conscious approach (e.g. employer support for commuting to work by bicycle or the development of a city centre incentive scheme, etc.).
- Adopting good practices in micromobility already tested.
- Strengthen public transport by consciously planning the uptake of micromobility.

Weaknesses:

- The measures to be introduced in urban transport conflict with the comfort of local residents.
- Age differences in the use of transport.
- Exposure to weather conditions, as reflected in the frequency of use of micromobility devices.
- The traveller's own comfort considerations.
- Fear of unfamiliar, new things.
- Rules for data collection, resulting from the use of smart devices.
- Lack of microgrid deployment is not conducive to further micromobility developments.

Opportunities:

- Social innovation.
- A move towards the idea of a walkable city, if combined with a rethinking of the use of cars in the city centre.
- Optimising the road network and parking.
- Freeing up working hours for residents not spent in traffic congestion.
- Expansion of digital, smart technologies, further development of the smart city.
- Moving towards a more resilient, green cityscape.
- Linking micromobility and public transport systems (common platform, common car sharing, etc.).

Threats:

- Both smart city transport solutions and the proliferation of micromobility devices could threaten the existence of certain jobs.
- The mass deployment of electric vehicles/vehicles/coaches will create greater network congestion.
- Increased energy dependence, economic downturn.
- Changing demands.
- Road accidents resulting from the flexible use of micromobility devices.
- Increasing battery "damage" may create aversion to the technology.

SCENARIOS

Sustainable and smart urban development measures can play an important role in keeping cities competitive and improving the lives of city dwellers. One of the areas where this can be achieved is through the transformation of urban transport. One way forward in the search for mobility is to start preparing the city and its population for the widest possible acceptance and use of innovative micromobility devices and digital and smart technologies. Here I have identified two possible scenarios.

The first one is that micromobility devices continue to spread rapidly, ahead of the development of smart urban transport strategies and their possible deployment. The old road networks will remain, cycle lanes will not be extended at a pace that will allow the micromobility passenger to be accommodated with their devices, congestion will continue to develop, and car congestion will continue. Pedestrian safety can now become dangerous from two directions. The untimely creation of intermodal hubs will increase tensions between car and alternative transport users in the daily travelling public, which in a post-covert period of constant crisis could further increase the mental fatigue and vulnerability of society. In such a scenario, disillusionment will prevail over time, the original development plan will be worn out, implementation will be delayed, the process may fail and the 'noisy' environment will not change.

In the second scenario, government actors, in partnership with the municipality, will continuously implement a sustainable urban development, with significant infrastructure investment supporting micromobility, along the lines of the smart urban transport strategy developed. An intermodal hub will be created, the sharing economy will be expanded, and the "upscaling" of micromobility from commuter/work to public transport will take place. The use of services and the interoperability

of systems will be facilitated by smart platforms that encourage individual activity, support recreation and health-consciousness. As a result, cities can be put on a development path to reduce the use of private cars in urban transport, promote alternative modes of transport, reduce health impacts, and move towards a more liveable and liveable urban vision (e.g. Oslo).

SUMMARY AND PROPOSALS FOR FURTHER RESEARCH

In this research, I examined the current state of the micromobility field, what trends can be identified and what influence they have. The European Union supports the introduction of more sustainable mobility solutions through which programmes and organisations, and how large cities in Hungary relate to innovative alternative means of transport.

The analysis has revealed megatrends and trends that influence the development directions of cities, such as continuous urbanisation (the interaction between technology and the city; the trend towards a boom in green actions; the organisation of cities into partnership networks), accelerating technological change and hyperconnectivity (the spread of wearable smart devices, data production; the impact of data analysis /personalised solutions/ and security risks; future mobility - the realisation of twin-transitions).), and the growth of consumption, with an increasing need for sustainable consumption.

Micromobility devices are present in the life of major cities, we experience them in our daily lives, and they are likely to continue their rise as a product of 'last mile solutions'. The emergence of shared mobility, with its potential to further shape the future landscape and infrastructure of cities through continuous technological innovation (and the spread of micro-networks). Alternative mobility devices will emerge in smart urban transport strategies to support public transport. At the same time, the public should be made aware of the potential of innovative mobility tools, rather than just car ownership, when they hear the word 'transport'. Older generations should be supported in acquiring the knowledge needed to use smart devices, as younger generations are more likely to take up technological innovations.

Research in the field of smart cities has intensified in recent years, both in Hungary and globally, and smart technology solutions have started to be integrated into sustainable urban development strategies in large cities, but the results so far have not yet led to spectacular breakthroughs. It would be good to involve as many city dwellers as possible in the data collection activities prior to development (crowdsourcing, citizen science), so that local needs can be channelled as accurately as possible. The process of building a sustainable and innovative mobility ecosystem in large cities can and should involve as many actors as possible, in order to provide a solid basis for the city's development in this direction. To this end, of course, we must not forget the basic premise that a smart and sustainable city can develop with smart and environmentally aware citizens, and that education is the way to achieve this.

REFERENCES

[1] 56/2017. (III. 20.) Korm. rendelet az egyes kormányrendeleteknek az „okos város”, „okos város módszertan” fogalom meghatározásával összefüggő módosításáról.
http://okosvaros.lechnerkozpont.hu/hu (Last viewed: 2024.09.15.)

[2] European Commission, The Megatrends Hub, 2024.
https://knowledge4policy.ec.europa.eu/foresight/tool/megatrends-hub_en
(Last viewed: 2024.08.27.)

[3] Gartner Hype Cycle, 2023.
https://www.gartner.com/ (Last viewed: 2024.08.27).

[4] M. Kelkar, R. Martins, G. Rodrigues, “A generational shift in urban mobility”, 2023.
https://www2.deloitte.com/us/en/insights/industry/public-sector/smart-mobility-traffic-solutions.html
(Last viewed: 2024.08.27.)

[5] H. Dediu, The Micromobility Definition, 2019.
https://micromobility.io/news/the-micromobility-definition
(Last viewed: 2024.08.27.)

[6] H. Dediu, The Micromobility Soulution, 2023.
https://www.youtube.com/watch?v=1RG0Ya07kWA&ab_channel=MicromobilityIndustries
(Last viewed: 2024.08.27.)

[7] J. Berényi, V. Hideg, D. Horváth, B. Remenyik, M. Siska, T. Fleischer, “A koronavírus-világjárvány és hatásai a közlekedési alágazatokban” In: A. Munkácsy – M. Jászberényi (szerk.): Fenntarthatóság és reziliencia a mobilitásban, Budapest: Akadémiai Kiadó., 2022.
https://mersz.hu/hivatkozas/m951feram_81_p4/#m951feram_81_p4
(Last viewed: 2024. 08. 27.)

[8] K. Heineke, B. Kloss, A. Mertens-von Rüden, T. Möller, C. Wiemuth “Shared mobility: Sustainable cities, shared destinies.” McKinsey Center for Future Mobility., 2023.
https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/shared-mobility-sustainable-cities-shared-destinies
(Last viewed: 2024.09.15.)

[9] Lancia Unveiled the Platum Series of E-Bikes as it Shifts Focus https://www.trendhunter.com/trends/platum-series, 2022.
(Last viewed: 2024.09.15.)

[10] D. Stewart, K. Ramachandran, “E-bikes merge into the fast lane”, 2022.
https://www2.deloitte.com/us/en/insights/industry/technology/smart-micromobility-e-bikes.html
(Last viewed: 2024.09.15.)

[11] Hogyan kezdődött a Segway története?
https://hu-hu.segway.com/about-the-brand
(Last viewed: 2024.09.15.)

[12] G. Michalkó, “A turizmus esszenciája” Budapest: Akadémiai Kiadó., 2023. https://doi.org/10.1556/9789634548683
(Last viewed: 2024. 09. 15.)

[13] M. Miskolczi, “Változó perspektívák napjaink turizmusában”, Budapest: Akadémiai Kiadó, 2023.
https://mersz.hu/hivatkozas/m1114vpnt_53_p1/#m1114vpnt_53_p1
(Last viewed: 2024. 09. 15.)

[14] B. Beszédes, Artificial Education Process Environment for Electro-Pneumatic System. In: IEEE - Óbudai Egyetem; IEEE (szerk.) 2023 IEEE 6th International Conference and Workshop Óbuda on Electrical and Power Engineering (CANDO-EPE): Proceedings. Danvers (MA), Amerikai Egyesült Államok: Institute of Electrical and Electronics Engineers (IEEE) (2023) 371 p. pp. 175-180. , 6 p.

[15] Micro-mobility startups worldwide, Statista.com., 2024.
https://www.statista.com/study/163775/micro-mobility-vehicle-startups-worldwide/
(Last viewed: 2024.09.15.)

[16] European Commission, Scalable Cities, 2022.
https://smart-cities-marketplace.ec.europa.eu/scalable-cities
(Last viewed: 2024.09.15.)

[17] European Commission, Sustainable and Smart Mobility Strategy, 2023.
https://transport.ec.europa.eu/transport-themes/mobility-strategy_en
(Last viewed: 2024.09.15.)

[18] European Commission, Sustainable urban mobility plans (SUMPs), 2023.
https://transport.ec.europa.eu/transport-themes/urban-transport/sustainable-urban-mobility-planning-and-monitoring_en
(Last viewed: 2024.09.15.)

[19] European Institute of Innovation and Technology (EIT), Urban Mobility, 2024.
https://www.eiturbanmobility.eu/
(Last viewed: 2024.09.15.)

[20] CIVITAS - Sustainable and smart mobility for all 2023.
https://civitas.eu/civinets/magyar-civinet
(Last viewed: 2024.09.15.)

[21] Sustainable Policy RespOnse to Urban mobility Transition, (Sprout), 2024.
https://sprout-civitas.eu/
(Last viewed: 2024.09.15.)

[22] Micromobility for Europe – Transforming urban mobility, 2024.
https://micromobilityforeurope.eu/
(Last viewed: 2024.09.15.)

[23] Shared Mobility Rocks – International Shared Mobility Conference, 2024.
https://www.shared-mobility.rocks/
(Last viewed: 2024.09.15.)

[24] Micromobility Europe, 2024.
https://micromobility.io/events/micromobility-europe
(Last viewed: 2024.09.15.)

[25] Li.me/blog, 2023.
https://www.li.me/blog/e-scooter-companies-want-regulation-heres-how-to-get-it-right
(Last viewed: 2024.09.15.)

[26] D. Szemere, T. Iványi, V. Surman, “Exploring Electric Scooter Regulations and User Perspectives: A Comprehensive Study in Hungary” CASE STUDIES ON TRANSPORT POLICY 15 Paper: 101135, 10 p. (2024)
https://doi.org/10.1016/j.cstp.2023.101135
(Last viewed: 2024.09.15.)

[27] Jovomobilitasa.hu, 2023.
https://www.jovomobilitasa.hu/80-100-ezer-roller-az-utakon-kinottuk-a-varosokat_-a-megoldas-a-mikromobilitas
(Last viewed: 2024.09.15.)

Analysis of Humanitarian Logistics Services during the Russian-Ukrainian War through the Exercise of the Hungarian Maltese Charity Service

Csilla Czeglédi
Faculty of International Management and Business
Budapest Business University
Budapest, Hungary
czegledi.csilla@uni-bge.hu

Ivett Gyarmati
Búzaszem Catholic Primary School and Primary Art School
Göd, Hungary
ivett.gyarmati@gmail.com

Ildikó Csuka Petőné
Alba Regia Faculty
Obuda University
Székesfehérvár, Hungary
csuka.ildiko@amk.uni-obuda.hu

***Abstract*— The study analyses the logistical processes related to the everyday work of the Hungarian Maltese Charity Service, the logistical system of humanitarian aid, the logistics of aid-related tasks, its difficulties, complexity and hardships during the Russian-Ukrainian war. The study also examines the changes that had to be introduced in day-to-day operations in order to make the logistical processes run smoothly and efficiently. The research methods used included secondary research, personal interviews and in-depth interviews. The study shows how the sudden refugee crisis, the reception, storage, distribution and distribution of large quantities of material donations, presented new logistical challenges to the staff and volunteers of the Hungarian Maltese Relief Service. It also examines the changes that had to be made to the logistics processes and analyses the effectiveness of the newly established systems.**

Keywords—logistics, logistics services, humanitarian action, humanitarian logistics, human resource)

I. INTRODUCTION

The Hungarian Maltese Relief Service carries out very important socio-economic activities in Hungary and in many foreign countries. After the outbreak of the Russian-Ukrainian war, the importance of the organisation's activities suddenly increased, as its objective is to help those in need and to respond to disasters. In carrying out these humanitarian activities, logistical processes are of particular importance. In order to understand humanitarian logistics, it is worth looking at the areas of activity and logistics practices of the Hungarian Maltese Relief Service.

In this article, we will therefore analyse the logistical processes related to the everyday work of the Hungarian Maltese Relief Service, the logistics of humanitarian aid, the logistics of relief work, its difficulties, complexity and hardships during the Russian-Ukrainian war. The authors also examine the changes that had to be introduced in day-to-day operations to make the logistics processes run smoothly and efficiently.

The choice of topic was influenced by the personal ties of one of the authors to Transcarpathia. The research was initiated based on the following questions. What new challenges did the refugee divide and the "logistics of people" in the disaster situation pose for the relief community? Did the sudden surge in donations pose a challenge in terms of storage, did it require expansion? What difficulties were encountered in delivering the donations? What improvements were needed? What changes in logistics did the crisis bring?

Most of the information and data needed for the analysis of the logistics processes was provided by the head of logistics of the Hungarian Maltese Relief Service in the framework of personal interviews. For the data collection on warehousing, individual face-to-face interviews gave an insight into the day-to-day work, and a longer conversation with the head of the central warehouse and the team leader of the warehouse staff gave a comprehensive picture of the usual tasks, as well as of the events that took place during the war and the challenges and changes that occurred as a consequence of the events.

As there is very little literature available in Hungarian on humanitarian logistics, the majority of the literature background was provided by English language publications and journals. In the literature review, the concept, purpose and principles of humanitarian logistics are introduced, followed by an analysis of the relationship between humanitarian logistics and disaster management. The impact of the Russian-Ukrainian war on the humanitarian logistics processes of the Hungarian Maltese Relief Service is explored. Following the daily events of the period after 24 February 2022, the difficulties are decribed and the solutions are outlined to the current problems. We are looking for the answer to the question of how the sudden refugee crisis, the reception, storage, distribution and distribution of large quantities of material donations, posed new logistical challenges for the staff and volunteers of the Hungarian Maltese Relief Service. What necessary changes had to be made in the logistics processes and how efficient the newly developed systems are.

II. HUNGARIAN MALTESE CHARITY SERVICE

In parallel with the expansion of its activities, by the mid-1990s the Charity Service had built up its national organisation. Seven regional organisations, including several smaller local groups, were set up under the national headquarters in Budapest (hereinafter referred to as the National Centre).

The regions operate under the direction and supervision of the National Centre in Budapest, but still independently, with their own HR and labour departments. In the 2010s, the Service's activities in disaster relief and rehabilitation have intensified both at national and international level. The red sludge disaster and flood relief in Hungary and the surrounding countries required an increasing number of volunteers and logistical planning. Accordingly, a separate group, the Emergency Response Task Force, was set up within the organisation to deal with disasters and emergencies.

It provides direct assistance to hundreds of thousands of people each year through its institutional services and the involvement of permanent and volunteer staff. The main categories - which are not entirely distinct, however, due to the complexity of the different services - are: donations; ambulance services, patient transport; model programmes; other social assistance programmes and projects; disaster management; foreign involvement; integration programme for refugees from war, of which the study will focus on disaster management and the integration programme for refugees from war.

A. Emergency management

The Charity Service carries out disaster management activities both in Hungary and abroad (typically in the countries of the Carpathian Basin). Its role is complex: it provides both health and civil protection services, as well as psychological and mental support to the victims. It is involved both in disaster response and immediate humanitarian assistance, but also in post-disaster rehabilitation. In Hungary, the organisation has already played a significant role in several natural disasters, such as the Red Sludge disaster, the floods of 2010 and 2013, and the 2016 ice storm and rainstorm. Abroad, he played a major role in relief efforts and fundraising during the 2014 Balkan floods, launched a relief shipment to crisis-hit Ukraine in the same year, and participated in the Be Drin European project for humanitarian and disaster relief in non-EU Balkan countries, a joint effort of several international organisations.

The Charity has a contractual public commitment to provide psychosocial support to the affected population and those involved in the rescue, as well as a commitment to provide health care and emergency relief and to build a temporary camp with a capacity of 200 people in the event of displacement.

B. Integration programme for refugees fleeing war

Following the outbreak of the war between Russia and Ukraine, the Maltese Relief Service has gradually launched an integration programme in all its regional centres from March 2022 to help Ukrainian refugee families. Refugees are provided with housing, employment, support for children's schooling, physical and mental health and integration into community life. In the programme, the organisation signs a cooperation agreement with the families for a period of one year. The one-year programme aims to prepare families to be able to manage independently in Hungary. The vast majority of families who have graduated from the programme have successfully become self-sufficient, with two-parent families often having built up some reserves before starting out. For mothers raising their children alone, the transition was more challenging, but most of them also moved on stronger after the programme ended. The integration programme is a very good opportunity for families wishing to settle down. For a full year they receive all the support they need to successfully start a new chapter in their lives. After this one year is over the Charity Service supports the families further in case of need [3].

Overall, it can be said that over the past decades, the Hungarian Maltese Relief Service has grown from a small, voluntary parish-based group to a national and even international aid organisation. Both the number of staff and the number of volunteers has grown exponentially, and the number of institutions it serves is constantly increasing. As a result, its organisations have become more fragmented and spread further apart, both geographically and in terms of their activities. As a result of major projects at national level over the last few years, it has been necessary to rethink and reform the logistics of the charity's operations in order to achieve greater efficiency. Before turning specifically to the logistical operation of the Charity Service, let us look at the theoretical background through a review of the literature on humanistic logistics.

III. HUMANITARIAN LOGISTICS

On the way to humanitarian logistics, in order to get a comprehensive picture of logistics processes, we need to understand the concept, purpose and principles of logistics. Fundamentally, logistics is an approach that requires a systemic mindset on the part of managers in order to achieve the best customer service. Logistics is therefore the efficient management of the movement and storage of materials. Accordingly, logistics covers processes such as transport, freight forwarding, fleet management, warehousing, inventory management, order fulfilment, logistics network planning, inventory management, supply and demand planning and logistics services [8].

The tasks of humanitarian logistics processes are the same as those of general logistics, complemented by the principles of humanitarian aid, which require aid organisations to be humanitarian, neutral and impartial.

A. Legal background and principles of humanitarian aid

The rules of international humanitarian law still in use today are based on the four Geneva Conventions adopted in 1949 and the Additional Protocols amending them.

The Geneva Conventions set out two basic principles for aid agencies: they require them to be humanitarian and impartial. The relevant provisions of the Conventions ensure the protection of the members of such aid organisations, which, unfortunately, is often violated. Among the other basic principles of humanitarian aid, neutrality must be highlighted. NGOs involved in aid must always be neutral, even if the aid is conditional on the consent of the opposing parties. These principles mean that all organisations must remain neutral in armed conflict, that aid must be distributed in accordance with the principles of impartiality and non-discrimination, and that only the degree of need should be a factor in access to aid. The principle of neutrality is sometimes interpreted differently, and some organisations consider that this principle and the fact that actions are carried out in a non-public manner are an obstacle to the effective protection of victims of conflicts. Neutrality in humanitarian action is of course different from the political neutrality that some states have adopted. Through humanitarian neutrality, the opposing parties accept that aid operations should not be considered either as an act of hostility or as a contribution to the war effort of one of the parties. Respect for the principle of independence is a fundamental requirement for all humanitarian organisations, which means that humanitarian action must not be subject to political, financial, religious, ideological or military pressure.

In a decision, the International Court of Justice in The Hague accepted that "the provision of assistance of a strictly humanitarian character to persons or forces in another country [...] cannot be considered a wrongful intervention if the aid is distributed without discrimination and without prejudice to its strictly humanitarian character. This is ultimately a good

definition and an acceptable limit to the right to provide humanitarian assistance." ([1], p.425) .

Humanitarian logistics is a process of planning, implementation and control that governs the efficient, cost-effective flow and storage of goods and materials, and the efficient delivery of related information, from point of origin to point of consumption, in order to alleviate the suffering of vulnerable people. This process encompasses a wide range of activities such as preparation, planning, procurement, goods movement, warehousing, tracking and tracing and customs clearance [6].

B. The relationship between humanitarian logistics and disaster management

.Humanitarian logistics encompasses several areas that can occur at any stage of emergency management. The type of disaster depends on the logistical needs of the situation. In the case of sudden-onset disasters such as earthquakes, hurricanes, terrorist acts, industrial accidents, a greater logistical effort will be required, at a higher cost due to time constraints. Slower-onset disasters, such as drought, famine, political refugee crises, have lower logistical demands at lower costs [4].

There are four phases in the disaster management cycle: prevention, preparedness, response and recovery [9]:.

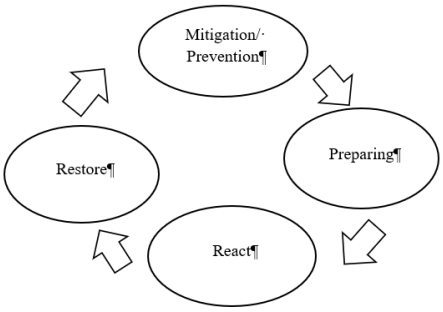


Fig. 1. Disaster management cycle flowchart [9].

The mitigation/prevention phase refers to legislation and mechanisms that reduce social vulnerability. This phase is the sole responsibility of governments and does not require the involvement of logisticians.

The preparedness phase covers the activities carried out before the disaster. In this phase, strategies are put in place for implementation, for a successful operational response. The aim is to avoid the most serious consequences of a disaster. This phase allows for implementation, and involves learning from and adapting to past disaster experiences to meet new challenges.

The response phase involves the implementation of various operations immediately after the disaster has occurred. This phase can be further broken down into two sub-categories:

a, The first is the "immediate response" group, during which the so-called "silent networks" or "temporary networks" are activated;

b, The second is the "recovery" group, which covers the restoration of essential services as soon as possible and, in this group, the delivery of rescue packages to as many people as possible.

In an emergency, cooperation and coordination between all parties involved is of paramount importance at this stage. Once contacted, all parties involved will try to respond as quickly as possible. The first 72 hours after a tragedy are crucial.

Recovery processes can be short and long term operations [2]:

- Short-term operations include fundraising activities, managing volunteers and donations, conducting damage assessments, and providing temporary shelter.

- Long-term operations can last for years after a disaster. Their aim is to help things get back to normal and improve the quality of life. This includes all activities, from providing medical care to rebuilding infrastructure, as well as humanitarian aid.

As we can see, there are many factors that influence the characteristics of humanitarian operations. Here we describe the two most commonly used: 'regular humanitarian logistics' and 'post-disaster humanitarian logistics'. Humanitarian logistics processes are very multifaceted and complex, requiring the effective cooperation of multiple support systems to operate properly. According to Holguín-Veras, logistics involves three complementary activities: a social network, a technical activities component and an underlying support system on which the other two depend. Since these factors are highly interdependent, a malfunctioning of any one of them can ruin the whole process. To be effective, all components must be examined and compared. The table below clearly shows the differences between the processes.

TABLE I. COMPARISON OF REGULAR AND POST-DISASTER HUMANITARIAN LOGISTICS

	Regular Humanitarian Logistics	**Post-Disaster Humanitarian Logistics**
Objective	Providing assistance to those in need	Alleviating human suffering as quickly as possible
Origin of Goods	Optional	Arrival of large quantities of goods in the disaster zone
Knowledge of Demand	Known	Unknown
Decision-making Structure	Structured interaction, controlled by a few decision-makers	Unstructured interaction, very many decision-makers
Frequency and Volume of Logistics Activities	Repetitive, constant flows	One-off, large and unexpected impulse
State of Social Networks and Connections	Relatively stable	Heavily influenced
State of Support Systems	Damaged but almost stable and functional	Influenced and dynamically changing

Own translation and editing [2]

All stakeholders involved in responding to humanitarian disasters operate according to a mission, responding to humanitarian needs. In order for an operation to be considered humanitarian, the principles of humanity, neutrality and impartiality must prevail [4].

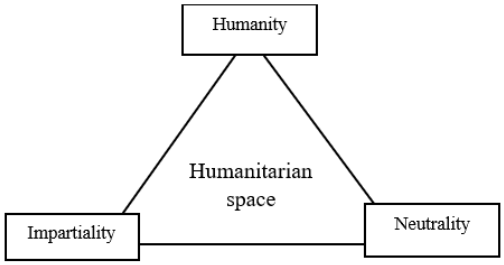


Fig. 2. Representing the context of humanitarian space- Maspero and Ittmann's model according to Jaharuddin [4]

The primary responsibility of all humanitarian organisations and other aid-related agencies is to operate in the context of the humanitarian space. There are six humanitarian aid organisations in Hungary, namely Catholic Charities, Hungarian Reformed Relief Service, Hungarian Maltese Relief Service, Ecumenical Relief Service, Baptist Relief Service and Hungarian Red Cross.

The world expects humanitarian service providers to meet the needs of people in distress during emergencies, disasters, conflicts and wars: to provide food, shelter, support for human rights, health and psychosocial care. To be effective, humanitarian organisations need to develop a long-term strategy that takes into account disaster trends and the vulnerability of the population. Long-term development and rehabilitation and coordination of the work of organisations are the strategic orientations that are the key factors in today's global humanitarian response. With environmental challenges and the changing dimension of security, the number of people and populations who are becoming increasingly vulnerable to disasters, armed conflicts and economic crisis is growing. One of the key tasks of humanitarian organisations is to make up-to-date decisions to meet changing needs and expand capacity [7].

C. The impact of war on logistical processes in the initial period of invasion

For a short time after the outbreak of the war, Hungarian-Ukrainian road transport practically ceased completely. In the first days, the aim was to interrupt ongoing transports and to leave the country quickly. At the border crossings, there were queues that lasted for days.

Humanitarian aid consignments began to be transported by road to the war-torn areas. Many countries in the European Union have exemptions for this type of transport. In Hungary, for example, trucks of at least EURO 3 class carrying humanitarian consignments on the initiative of a humanitarian organisation are exempt from the weekend and holiday ban, and the NAV has fast-tracked the removal of aid consignments from Hungary to Ukraine. Slovenia, the Czech Republic and Germany have exempted lorries carrying humanitarian consignments by road from road tolls, while Poland has temporarily derogated from the rules on driving and rest periods for drivers (not only for humanitarian consignments).

Rail transport has also become impossible in the Russia-Ukraine relationship. Transit traffic has also come to a complete standstill, including container traffic between China and Záhony, which was in Russian transit. Russian troops marching in also physically destroyed the connection between the two countries. Traffic was diverted to new routes. The diversion of goods to the Ukrainian port placed a greater burden on the operating stations of other countries. A period of turmoil began. Manufacturers and distributors focused on exploring new markets.

In the container transport market, the transport of non-essential cargo to and from Russia was suspended, with the exception of foodstuffs, medical supplies and humanitarian supplies. Road haulage and freight forwarding companies have also seen their costs rise significantly as a result of record fuel prices.

All areas of logistics services have been negatively affected by the outbreak of war, not only in neighbouring countries, but also worldwide.

IV. HUMANITARIAN LOGISTICS ACTIVITIES OF THE HUNGARIAN MALTESE RELIEF SERVICE DURING THE WAR

A. Research objective

The main objective of this research is to explore the impact of a disaster situation on humanitarian logistic processes through the study of the activities of the Maltese Relief Service. The research also aims to describe the process of humanitarian assistance in crisis situations and to observe the extent to which the logistical activities are affected by the invasion. The guiding principle of our research is therefore to analyse humanitarian logistics by looking at practical experiences. After delving deeper into the topic, we first looked at the events related to humanitarian assistance after a disaster within the research questions. In what respects did the sudden refugee crisis and the provision of supplies to the large numbers of refugees encounter difficulties and how were the problems solved? What aspects of this process were followed? We sought to find out how much the massive donations have complicated the life of the charity. Were there any difficulties in accommodating the donations, did any innovations or extensions have to be made at the storage level? In addition to the creation of staff and material conditions, the challenges encountered in the process of delivering the goods are also analysed.

B. Research method

Two research methods were used:

1. secondary sourceresearch based on internal documents, annual reports, case studies, and

2. primary research in the form of in-depth interviews.

The qualitative research method was used to collect data. Qualitative techniques are an excellent tool to use when we want to find answers to our hypotheses through a small sample, in this case the practices of a specific organisation. In-depth interviews also have the advantage of being narrative, where the interviewee is free to talk about the key events in his or her activities, giving the respondent greater freedom to express his or her thoughts. Among the tools provided by the guide, the funnel method [5] is used in the structure of the in-depth interview, which starts with a general discussion of the charitable activities of the Charity, followed by a discussion of the events of the period during the war and the challenges and responses to the invasion. We then explore the improvements that were needed and conclude with an evaluation of the innovations and the lessons learned.

Our interviewees were the Head of Logistics at the Hungarian Maltese Charity Service, who answered questions in a 90-minute interview. In addition, she provided me with several data in the form of internal documents, which contributed greatly to the understanding of the research area and the exploration of the data. In addition, field visits and personal interviews with the warehouse manager and the team leader were also provided. They gave a tour of the warehouse of the logistics centre, showed the warehouse area, explained in detail the stocking techniques and guidelines used to organise the product groups. All the while, they readily answered the questions that had been prepared in advance. The field visit was very helpful in understanding the warehousing processes.

C. Humanitarian service at the start of the war

The Hungarian Maltese Charity Service also worked in partnership with other members of the Charity Council to care for refugees. The Charity Council divided the border crossings between the organisations, so that the Hungarian Maltese Charity Service received refugees from Ukraine at Beregsurány. The Charity Council allocated 500 million forints to all its members, including the Maltese Relief Service, to take care of refugee-related tasks. At the Beregsurany help point, refugees were welcomed with hot tea, food and first aid. Donations were stored in buildings donated by the municipality, and the school hall was converted into temporary accommodation with camp beds. An office and accommodation for helpers were set up in the parish hall. The physical environment was set up within days, while psychosocial support was also provided from the start.

D. Challenges in the logistics centre

After the outbreak of war, the central warehouse in Budapest was open seven days a week to receive donations. The first four to six weeks were mostly spent receiving donations and only a part of them could be sorted and delivered to their place of use. During this period, 179 volunteers assisted the full-time staff, who worked from dawn to dusk seven days a week. The main tasks were receiving, sorting and transhipping donations. With the help of the volunteers, small food parcels were prepared for the families that arrived.

Due to the volume of incoming donations, the warehouse was stretched to its limits. The nature of the donations ranged across a wide spectrum: durable food, hygiene products, basic medicines, bandages, flashlights, batteries, mattresses, blankets, camp beds were received based on donations from residents and businesses. Difficulties were encountered in accommodating this large and varied range of products. In order to increase the capacity, an event tent of nearly 300 m2 was set up in the courtyard to accommodate donations that could no longer fit in the warehouse. A temporary warehouse was set up in Budaörs and Waberer's - WSZL also offered its storage capacity, where several pallets of goods could be transferred.

As time went on, the volume of donations decreased significantly, but still at a much higher rate than in similar months in previous years. In the long term, the increased volume of donations has raised the idea of the need for warehouse expansion. In the spring of 2024, construction of a new warehouse building will start, which, once completed, will allow for a more specialised storage of the various products. In addition to the difficulties with warehouse capacity, there was also a need to increase stock levels in the material handling area in order to move materials more quickly and efficiently. To handle the large volumes of loads that need to be moved at the same time, new equipment was acquired: 2 new forklift trucks and a pallet wrapping machine were added in 2022. As there were simultaneous releases or receipts of donations at several locations, the multiple forklifts allowed for parallel and continuous work, which was essential to optimise logistics processes. Furthermore, the increased demand resulted in the need to keep a larger safety stock.

E. The Logistics of Humanitarian Aid

The donations were delivered to the local warehouse of the Beregszász Maltese Charity through the Beregsurány border crossing using 3.5-ton vans. It became clear after the first few deliveries that this mode of transportation would not be optimal from a logistical standpoint. Delivering the shipments required reorganization. To ensure more efficient transportation of larger quantities of donations, the Charity enlisted the help of an external transportation company, and starting from the end of March 2022, they began using articulated trucks to transport donations from the central warehouse in Budapest to Beregszász via Záhony. In the meantime, they continued deliveries with their own vans, with 85-90% of the donations reaching Beregszász. Additionally, several shipments were also sent to Mukachevo, Uzhhorod, Kerekespuszta, and Ivano-Frankivsk. Logistically, there was a need to establish an intermediate warehouse near the border, which was made possible by the offer of an apple storage facility in Tarpai. This warehouse served as a transshipment point for those aid shipments that were routed through transloading to reach their final destination.

The Charity's fleet of vans crossed the border 129 times in 2022, while Waberer’s - WSZL articulated trucks made 29 crossings, delivering a total of 500 tons of humanitarian aid to Ukraine. The partnership with Waberer’s - WSZL began in 2019 but reached a new level in 2022 when they took on the delivery of aid shipments to Ukraine. They also purchased two new vans and renewed their truck fleet with a 15-ton, 2-year-old Mercedes Atego and a brand-new 7.5-ton DAF truck. In addition, a parking area was established next to the warehouse for the renewed fleet. This updated fleet significantly increased efficiency: while a 3.5-ton small truck can carry 4 or 5 pallets, the larger truck can hold up to 18. In contrast, the WSZL articulated truck can transport up to 33 pallets. The development of this logistical partnership was a huge step forward in the distribution of donations [9].

.

For humanitarian shipments entering Ukraine, a simplified customs procedure was introduced, requiring the detailed breakdown of items to be listed by product categories. This greatly facilitated the preparation of documentation for customs clearance and expedited the border crossing process.

Our research indicates, based on interviews and reports related to logistical processes, that after establishing the necessary physical conditions, the first significant difficulty arose in the logistics of personnel. The daily influx of thousands of refugees created a very chaotic situation at the border crossing, necessitating the rapid establishment of a clear system.Within days, the Charity's staff developed a transparent and efficient process. At the moment of the first encounter, they established the primary and most important humanitarian condition: "to alleviate human suffering."

The subsequent phase of the process was to evaluate the requirements. Once the needs had been identified, targeted assistance could be provided for the subsequent stages. It is evident that the logistical processes commence with a needs assessment, which is a crucial element in the endeavour to guarantee that the assistance furnished to arrivals is as targeted as possible.It is our contention that the concept developed in this phase was transparent and functioned in a highly organised and efficient manner. The organisation of assistance providers in proximity to the refugees was conducted in a systematic manner. However, in order to guarantee the optimal functioning of these processes, it was necessary to hire new personnel, which resulted in an expansion of the charity's workforce.

The difficulties that arose in the logistics centre will be introduced in the following. In accordance with the unanimous opinion of the leaders who were interviewed during in-depth discussions, the greatest challenge faced by the logistics centre in 2022 was related to tasks that arose as a consequence of the Russian-Ukrainian war. The sudden influx of substantial quantities of donations presented considerable challenges with regard to their reception, sorting, and storage. Despite the assistance provided by volunteers, the staff were still working at a considerable pace. It became evident within the initial period that, due to the limitations of the available space, it would not be feasible to store the received products in an appropriate manner within the existing 5000 m² area. In order to accommodate the donations, the site was expanded and external storage was utilised. It became evident that the establishment of a new warehouse building would be essential in the long term. The necessity for expansion had been previously contemplated, but the urgency of the situation became apparent at this juncture.

V. SUMMARY

This study examined the influence of the Russian-Ukrainian conflict on the humanitarian logistical operations of the Hungarian Maltese Charity. The research methods employed included the utilisation of secondary sources, personal interviews and in-depth discussions. In the aftermath of the conflict's inception, the Hungarian Maltese Charity was confronted with the formidable task of managing the considerable influx of material donations. It became evident that an increase in storage capacity was necessary during the warehousing process. This was addressed by setting up tents in the courtyard and utilising external locations. Furthermore, the necessity arose for the acquisition of additional equipment with the objective of optimising the efficiency of material handling operations. The storage system was reformed and a new barcode system was introduced with the objective of achieving more transparent and traceable inventory management. Moreover, to guarantee prompt and efficacious assistance in the event of an unforeseen catastrophe, the safety stock was also augmented. These developments have had a beneficial impact on the present situation, facilitating the daily humanitarian supply with the new innovations.

A significant challenge was the lack of suitable equipment for the transportation of large quantities of shipments. This issue was subsequently addressed through the utilisation of articulated trucks by Waberer's-WSZL for the transportation of donations. In addition to providing transport vehicles, the company also assigned a dedicated staff member to oversee the deliveries to Ukraine. This logistical partnership also facilitated the provision of supplies within the domestic and regional contexts. From the perspective of ensuring an adequate supply and the optimal division of labour, it was also necessary to expand the workforce. The implementation of local outsourcing resulted in more efficient processes, necessitating a reduced level of human input. Another consequence of the disaster was the acquisition of practical knowledge in the field of humanitarian activities. The humanitarian (logistical) work conducted in response to the war disaster enabled the Charity's staff to develop the capacity to resolve issues that arise in crisis situations, thereby providing a foundation for future endeavours.

REFERENCES

[1] Blahó A.–Prandler Á. (2016): International organisations and institutions.Budapest: Akadémiai Kiadó

[2] Holguín-Veras, J. – Jaller, M. – Van Wassenhove, L. N. – Pérez, N. – Wachtendorf, T. (2012): On the unique features of post-disaster humanitarian logistics. Journal of Operations Management, Vol. 30 Iss: 7-8, p. 494-506.

[3] Hungarian Maltese Charity Service: Annual accounts 2022, Service log 2022.02.24-2023.02.24

[4] Jaharuddin, N.S., I. Saadiatul (2014): Management Science A Collection of Readings, p 6-16.

[5] Naresh K. Malhotra– with Simon J. (2017): Marketing research. Budapest: Akadémiai Kiadó.

[6] Overstreet, R. E. – Hall, D. – Hanna, J.B. – Rainer, R. K. Jr, (2011): Research in humanitarian logistics. Journal of Humanitarian Logistics and Supply Chain Management, Vol. 1 Iss 2, p. 114-131.

[7] Sáfár B. (2018): Theoretical and practical issues of humanitarian aid in the light of the international activities of the Red Cross, Budapest, PhD Dissertation

[8] Szász L.–Demeter K. (2017): Supply chain management. Budapest: Akadémiai Kiadó

[9] Van Wassenhove, L. N. – Pedraza-Martinez, A. J. (2010): Using OR to adapt supply chain management best practices to humanitarian logistics. International Transactions in Operational Research, Vol. 19 Iss: 1-2, p. 307-322

Experiential learning implementation opportunity in engineering education

Mónika Seres
Alba Regia Faculty
Obuda University
Szekesfehervar, Hungary
seres.monika@amk.uni-obuda.hu
ORCID: 0009-0003-1288-3111

Monika Pogatsnik
Alba Regia Faculty
Obuda University
Szekesfehervar, Hungary
pogatsnik.monika@amk.uni-obuda.hu
ORCID: 0000-0002-2698-7291

***Abstract*— The recent changes of the past few years have posed significant challenges for students, lecturers, and mentors involved in dual education programs. In addition to events that negatively impacted education, the rapid advancement of artificial intelligence — now increasingly integrated into everyday life — has necessitated the swift introduction of new learning and teaching strategies. The growing importance of adaptability, self-awareness, and flexibility in students' approaches to learning, and later, to their professional careers, is becoming ever more apparent. At Obuda University, it has become imperative to develop a curriculum that aligns more closely with these evolving circumstances, thereby enhancing students' chances of success throughout their academic journey. To this end, the introduction of courses in learning methodology and tutoring was essential, as they effectively support the development of key competencies such as efficient learning strategies, motivation, and self-awareness. Several key considerations emerged in the transformation of university education: the need for dynamic and continuously evolving professional programs that can better respond to societal demands and challenges; an emphasis on experience-based, engaging learning environments; the achievement of improved academic outcomes; the development of methods that foster the foundation for lifelong learning; and the promotion of fairness and equity within the educational system.**

Keywords*— *experiential learning, personalized professional development, student-centeredness, sustainability, change of attitude, self-awareness, methodological renewal, engineering pedagogy

I. Introduction

The Covid-19 pandemic in 2020 rapidly disrupted the structure of secondary and higher education, with many schools and individuals struggling to adapt to the sudden shift. Teachers, students, and parents alike faced unfamiliar challenges, as they had never encountered such a situation. Some quickly adapted by utilizing online platforms to simulate traditional classroom environments for face-to-face instruction. In many instances, however, educators resorted to simply uploading teaching materials to various online platforms, leaving students to independently process the content at home. This exposed a widespread lack of preparation for online education, both in terms of technical infrastructure and pedagogical methods [1].

Following the pandemic, there has been a growing demand for the development and integration of new teaching methods in university curricula to improve the quality and efficiency of e-learning. These methods include blended learning, research-oriented learning, project-based learning, and experiential pedagogy. This demand is not solely driven by the pandemic's push for rapid e-learning adoption but also by the need for educational systems to respond swiftly to ongoing industrial, technological, and societal transformations [2].

In our study, we conduct a secondary literature review to explore how the education system can adapt and renew itself in response to the accelerating socio-economic and technological environment. Additionally, we incorporate data from primary research based on student feedback. Higher education is currently seeking to align its curricula with the evolving demands of Industry 4.0. To achieve this, flexible programs and course content are essential for enabling personalized professional development. Ensuring a solid foundation of knowledge is crucial for fostering sustainability and innovation in education, while solidarity is also gaining prominence as a key aspect of the educational process [3].

The primary aim of this research is to demonstrate the relevance and necessity of teaching learning methodologies and implementing experiential pedagogy in contemporary higher education. A further objective is to provide educators with more effective tools for selecting pedagogical methods and, where appropriate, introducing entirely new approaches to better support today's students in becoming successful engineers.

II. Stages in the Development of Engineering Education

After the first industrial revolution, engineering education started in universities in many countries, including Hungary. The development of production, infrastructure and military technology necessitated the creation and continuous development of engineering education. As a result of the second industrial revolution, a continuous search for a balance between the theoretical and practical aspects of engineering appears, and the development of modern engineering education is also connected to this. To perceive technology, art and craft as a global unit; to create chemical and electrical engineering as an independent discipline; and new concepts are incorporated into university engineering education (1880-1940).

Between the 1950s and the 1980s, following the third, i.e., digital industrial revolution, the first programs appeared in some contemporary engineering disciplines, including: biomedical engineering, electronics, computer technology, robotics and mechatronics. The founding of the ERASMUS program in 1987 plays an important role not only in domestic but also in international engineering education [4]. This is

when the transition to more modern, student-centered educational strategies begins in education. The integration of information technologies into education and management, the wide-ranging design of laboratory and research practices, the transition from analog to digital, and the implantation of computer-supported devices are also launched together with these learning strategies. Quality management systems are also integrated [5].

III. Current engineering education

The turn of the millennium brought another important change in engineering education. There is a clear emphasis on student-centered education instead of the traditional teacher-centered approach [6]. The popularity of active learning methodologies, experiences and environments has facilitated the incorporation of Industry 4.0 technologies and methods into engineering programs worldwide. Topics such as cloud-based computing, simulation methods, autonomous robots, and additive manufacturing have been researched at universities since at least the 1990s. Today, these technologies and methods are widely used in most engineering programs at all levels. The use of e- learning methods, e-portfolios and virtual laboratories are coming to the fore in engineering education, and last but not least, the growing importance of internationalization is also appearing in higher education. There is also a need for the democratization of education and more and more lectures and courses will be publicly available to everyone on the Internet [7].

IV. Project pedagogy

The appearance of project pedagogy in higher education presented teachers and students with such challenges that, abandoning teaching methods based on traditional principles, it led university courses to a new type of educational form that offers support and learning styles [8]. The instructor is present as a coordinating actor in a continuous advisory role, forming a partnership with the students. This requires the cooperation of teacher and student, individual efforts, preparations, knowledge sharing, thinking together, cooperation and the division of work within the team. So, you need to know the individual software skills, the basics of communication, the basic rules of cooperation in a group. During joint work, the instructor and the students must learn to cooperate as partners, which is a great challenge for both parties, since this requires completely different skills than the traditional, frontal form of education. With this form of learning, it is possible to create active knowledge instead of passive knowledge, so that they can acquire additional skills in order to solve the problems that arise during the work well and quickly, in addition to the learning processes.

Currently, an important feature of engineering education should be preparation for lifelong learning. Students need methods that can be integrated into their everyday work and work automatically during specific learning, are present in their everyday life without being noticed and learn playfully with their help. The knowledge acquired in higher education helps to get a good job in the workplace, and practice helps to deepen knowledge and acquire additional knowledge. The success that can be achieved in the labor market therefore largely depends on the university education and the preparation for lifelong learning. Self-directed learning is one of the cornerstones of lifelong learning. Today 's employees must know how to take control of their learning processes, they must be able to plan, develop, adapt and change in today's digital, interactive and global society [9].

V. The appearance of artifactal intelligence in unuversity education

Students are increasingly using artificial intelligence on the Internet, in addition to various search software, and are calling on artificial intelligence to help them learn and prepare for exams. The greatest achievement of the 21st century so far is the development of artificial intelligence. Its spread has brought about many changes in everyday use. Its appearance in education has already caused quite a few pleasant surprises, but also many disappointments [10] However, students use it more and more frequently and better in studying, writing assignments or preparing for exams. That is why education should not ignore the opportunities provided by AI, which can also appear as motivating and inspiring factors for students to learn. Based on previous experience, instructors can use it well when compiling educational materials, organizing and expanding knowledge, and in connection with the compilation and evaluation of tasks. And the students when sketching, summarizing, extracting, explaining and answering specific questions. The use of AI still raises many questions for universities, each institution decides on its use independently for the time being, but it is already clear that its use in education cannot be ignored. Those who fall behind are left out.

VI. Experience pedagogy

In the course of the development of the history of education, such pedagogues and the methods they formulated played a decisive role, which provide the foundations of experiential pedagogy in today's sense. Such is Dewey's creed from 1897, according to which school is life itself. Dewey believed that students are always guided by some individual goal, the realization of their own interests, during social activities, so he tried to find the appropriate tools and develop them further. In his opinion, life, society, and the process of education provide the goals of education and provide the necessary means for their realization [11].

Engineering students may encounter many challenges during their studies. The biggest challenge is learning the subject of mathematics. The teaching of mathematics plays a prominent role in all engineering courses, and the understanding and application of analysis, linear algebra, and differential equations are a challenge for many students. Time and resource management plays an important role in the formation of students' learning habits. Managing their time and creating a good time plan not only plays an important role in learning, but is also an important factor in work, project tasks and in their personal lives. With the introduction of the learning methodology subject, the legitimate expectation that students can prepare for proper time management is realized.

Within the framework of the subject, they get a complete picture of the importance and methods of time management.

Industry 5.0 expects a multidisciplinary approach from young engineers, as the changes currently taking place are increasingly forcing those involved in engineering work to be proficient in other disciplines as well. They must be able to connect their knowledge in both the natural sciences and the IT sector. They must acquire the ability and skills to adapt technological and innovative changes as quickly as possible. It is increasingly necessary to work in a team, which requires effective communication from those involved in the projects. For this, it is essential to get to know the communication methods as widely as possible during the training, and learn them playfully. This communication can be realized during the subject and tutoring in the current training [12].

The aforementioned challenges make it almost mandatory for them to need continuous personal and professional development, to learn new methods, and to respond to changes as quickly as possible. This requires assistance during engineering training. In order to effectively manage the challenges of time and resource management, several methods and techniques are available, which they can master at the beginning of their studies and can continuously develop and customize. By the time they face the challenges of the labor market, they can already operate time management at a skill level [13].

They get a complete picture for creating the task list and the time estimate, with which they can allocate their time efficiently, making priorities transparent. By using the calendar function, you can avoid wasting time and focus on studying and working. By preparing weekly and monthly task plans, it is easier to tune in to the task, the tasks become more transparent, and it is also verifiable. In this way, control is integrated into their everyday life, as it is clearly visible which tasks they can and cannot complete during this period. With the help of resource auditing, they can accurately assess the available resources, it is possible to replace the missing elements in time, so project management can be implemented more efficiently. It provides additional assistance in the efficient distribution of resources and in decision-making. If the subject education and practical application of these are well established, it represents such positive progress in the students' lives that it can function automatically in the future as well. By acquiring positive experiences, the method of experiential pedagogy is realized.

Experiential pedagogy therefore connects experience, learning and preparation for life. Learning is based on one's own experience and the acquisition of experience, which can be easily achieved by mastering time management.

During the teaching-learning process, students acquire knowledge and professional knowledge, thereby expanding their knowledge base and proficiency. In this process, the instructor gives a lecture, gives explanations, helps in the processing of the course material by making sketches and presenting learning methods. The knowledge transferred in this way must become the students' own knowledge, the internalization process must take place [14].

The experience itself is a series of events that affect both the instructor and the students. It involves and activates the participants in joint work. Solving tasks and acquiring knowledge becomes a pleasure, the lesson takes place between fun and memorable processes, and learning turns into a source of joy. This is a motivational training that provides the teacher with additional opportunities to transfer knowledge that was otherwise difficult to achieve during face-to-face teaching.

The purpose of experience-based education is therefore to attract attention, to keep goals in mind, to build community, to increase student-centeredness and interactions [15]. The expectation is fulfilled that the student can identify with the course material, his interest increases, and his motivation remains continuous. The students' communication skills develop, they dare to speak about the topic, come up with ideas, and actively participate in class work.

Based on the survey conducted during the teaching of the learning methodology subject, it can be said that the students solved the tasks with pleasure and incorporated the experiences into their everyday lives. Their impact is long-term, as the experiences remain and can continuously inspire. With the help of experiential pedagogy, a good communication channel can be established between the instructor and the students, all of which is of great help in the teaching of later subjects. The teacher-student relationship formed in this way can be continuously built upon, planned, and concepts developed.

Nowadays, it is necessary to talk about the introduction of the 5.0 education strategy for engineering education and the training courses should be restructured accordingly. University education should be reorganized in such a way that 16 key features can be reflected in education [16]. Engineering education adapted to Industry 5.0 is based on well-proven educational models, but at the same time it must incorporate forward-looking, innovative aspects and rely on advanced technologies to effectively and successfully face global social and environmental challenges. International experiences should not be ignored during innovation, the positive feedback of dual training should be incorporated in the development of new strategies. It is essential to coordinate the opportunities provided by the labor market, revolutionary innovations, artificial intelligence, its appearance in education and test new methods based on these [17], [18], [19].

VII. Summary

In our study, we presented the findings of a secondary analysis based on data from primary research conducted during the instruction of the "Learning Methodology" course, which is part of the newly implemented curriculum at the Alba Regia Faculty of Obuda University. The analysis also incorporated the experiences gained through this process. Our research addressed the development of engineering education, the current state of engineering education, and the growing need to integrate project-based and experiential learning approaches.

Based on student feedback, it was observed that the practical application of time management skills not only enhances their academic performance but also contributes positively to their daily lives. Time management proves to be an essential tool for balancing academic and work-related

responsibilities. Furthermore, our study explored the anticipated changes in engineering education in relation to Industry 5.0, highlighting the increasing importance of soft skills. These skills, particularly problem-solving, adaptability, and teamwork, are becoming crucial for students to develop at an advanced level.

1. figure Experiential pedagogy - word cloud

Students entering higher education tend to retain the study habits acquired in secondary school, often struggling to adapt to the more demanding pace of university education. It is therefore essential to support students in establishing effective time management, developing intrinsic motivation, and selecting appropriate learning strategies. We believe that by incorporating experiential learning and future-oriented teaching methods, we can lay a strong foundation for experiential education, ultimately facilitating the highest level of knowledge acquisition, both in subject matter and professional expertise.

REFERENCES

[1] A. Aristovnik, K. Karampelas, L. Umek, and D. Ravšelj, "Impact of the COVID-19 pandemic on online learning in higher education: a bibliometric analysis," in Frontiers in Education, vol. 8, p. 1225834, Aug. 2023, Frontiers Media SA.

[2] N. A. Bizami, Z. Tasir, and S. N. Kew, "Innovative pedagogical principles and technological tools capabilities for immersive blended learning: a systematic literature review," Education and Information Technologies, vol. 28, no. 2, pp. 1373-1425, 2023.

[3] A. T. Gorski, E. D. Ranf, D. Badea, E. E. Halmaghi, and H. Gorski, "Education for sustainability—Some bibliometric insights," Sustainability, vol. 15, no. 20, p. 14916, 2023.

[4] G. Croce and E. Ghignoni, "The multifaceted impact of Erasmus Programme on the school-to-work transition: a matching sensitivity analysis," Research in Higher Education, pp. 1-23, 2024.

[5] M. S. Assingkily and N. Putri, "An integrated system for learning quality management based on the Tridharma of higher education," EDUMALSYS Journal of Research in Education Management, vol. 1, no. 1, pp. 10-21, 2023.

[6] K. H. D. Tang, "Student-centered approach in teaching and learning: what does it really mean?," Acta Pedagogia Asiana, vol. 2, no. 2, pp. 72-83, 2023.

[7] M. T. Judit and P. Monika, "A Comprehensive Study on Cybersecurity Awareness: Adaptation and Validation of a Questionnaire in Hungarian Higher Technical Education," ACTA POLYTECHNICA HUNGARICA, vol. 21, no. 10, pp. 533–552, 2024.

[8] H. Hou, J. H. Lai, and H. Wu, "Project-based learning and pedagogies for virtual reality-aided green building education: case study on a university course," International Journal of Sustainability in Higher Education, vol. 24, no. 6, pp. 1308-1327, 2023.

[9] W. C. Brandt, "Measuring student success skills: a review of the literature on self-directed learning," 21st Century Success Skills, National Center for the Improvement of Educational Assessment, 2020.

[10] F. Kamalov, D. Santandreu Calonge, and I. Gurrib, "New era of artificial intelligence in education: towards a sustainable multifaceted revolution," Sustainability, vol. 15, no. 16, p. 12451, 2023.

[11] J. Frank, Teaching in the Now: John Dewey on the Educational Present. West Lafayette, IN: Purdue University Press, 2019.

[12] M. T. Judit and P. Monika, "The Presence of Cybersecurity Competencies in the Engineering Education of Generation Z,"Acta Polytechnica Hungarica, vol. 21, no. 6, pp. 107–127, 2024.

[13] D. L. Stone and D. L. Deadrick, "Challenges and opportunities affecting the future of human resource management," Human Resource Management Review, vol. 25, no. 2, pp. 139-145, 2015.

[14] S. B. Rawal, "Exploring exciting ways to teach: a comprehensive analysis of modern teaching methods," The Journal of Social Sciences Studies and Research, vol. 4, no. 03, pp. 105-110, 2024.

[15] T. Muniandy and N. Abdullah, "A comprehensive review: an innovative pedagogy for future education," International Journal of Online Pedagogy and Course Design (IJOPCD), vol. 13, no. 1, pp. 1-15, 2023.

[16] Lantada, A. D. (2020). Engineering education 5.0: Continuously evolving engineering education. International Journal of Engineering Education, 36(6), 1814-1832.

[17] Gy. Györök, B. Beszédes. Fault-tolerant Software Solutions in Microcontroller Based Systems. In: Orosz, Gábor Tamás (szerk.) AIS 2017 - 12th International Symposium on Applied Informatics and Related Areas organized in the frame of Hungarian Science Festival 2017 by Óbuda University: Proceedings. Székesfehérvár, Magyarország: Óbudai Egyetem (2017) 204 p. pp. 7-12., 6 p.

[18] Gy. Györök, B. Beszédes. Duplicated Control Unit Based Embedded Fault-masking Systems. In: IEEE - Szakál, Anikó; IEEE (szerk.) 2017 IEEE 15th International Symposium on Intelligent Systems and Informatics (SISY): 15th IEEE International Symposium on Intelligent Systems and Informatics. New York, Amerikai Egyesült Államok: IEEE (2017) pp. 283-288. Paper: 8080567, 6 p.

[19] Gy. Györök, B. Beszédes. Fault tolerant power supply systems. In: Orosz, Gábor Tamás (szerk.) 11th International Symposium on Applied Informatics and Related Areas (AIS 2016). Székesfehérvár, Magyarország: Óbudai Egyetem (2016) pp. 68-73. , 6 p.

Vocational education and tarining teachers' views on climate change

János Zoltán Trenka
Óbuda University
Alba Regia Faculty
Székesfehérvár 8000
E-mail: trenka.zoltan@uni-obuda.hu

Monika Pogatsnik
Óbuda University
Alba Regia Faculty
Szekesfehervar, 8000
pogatsnik.monika@amk.uni-obuda.hu

***Abstract*— Today, climate change and its potential impacts are receiving increasing attention. We only have to think of the recent drought summers and warming weather, or the large amounts of rainfall that have fallen in a short period of time. We also know that we need to be prepared for these changes. One possible way of preparing for these changes could be to address them in the field of education. That is why it is important to find out what those involved in education think about climate change and how prepared they feel. In this article I summarise the results of 33 responses from two vocational training institutions.**

Keywords – education, climate change, security risks

I. INTRODUCTION

The fact of climate change and its negative effects is no longer a matter for debate. We are seeing events in our environment, drought conditions, large amounts of precipitation in a short period of time, which prove that these processes are already happening and are with us. These processes are already being investigated by science and have been proven in numerous studies. Examples include the study by Qureshi Haris Uddin; Shah Syed Muzzamil Hussain, who examined temperature data from Pakistani cities. In their study, they show that rising annual and seasonal temperatures over urban areas of the country can lead to increased health risks, heat wave events, and high daily household and municipal water demand due to the development of low pressure zones [1]. Evidence of rising temperatures can be traced over long time periods in studies. An example of this can be seen in the book Climate and Society in the Eurozone, which shows more than 200 years of temperature data for the city of Basel [2].

Climate change impacts not only affect people, but also surface waters, for example. In a study published in 2012, Philipp Schneider; Simon J. Hook used satellite data to investigate how surface water temperatures have changed. The results of the study show that there have been significant warming trends between 1985 and 2009. The most rapid warming was observed in northern Europe [3].

We see that climate change is affecting all aspects of our lives and will continue to do so in the future. Future generations must be prepared for these impacts and weather anomalies. One way to prepare for this is to make efforts to do so more prominent in education. Education should equip the future generation with competences such as teamwork, participation in a sustainable economy, or creativity [4], all with a view to adapting to climate change.

The questionnaire survey on which this article is based was based on reading an article on the Hungarian perception of security, in which a representative survey showed that climate change is one of the most important factors affecting security. This is followed by uncontrolled migration. [5] After reading the article, I was curious to know what my colleagues involved in teaching in the field thought about the topic. In the questions that followed, I focused not only on thoughts related to climate change, but also tried to explore how they saw climate change and adaptation as something that could be incorporated into education. The starting point for the questions on education was provided by writings such as the article by Oversby, J. [6], or the paper by Stevenson, R. B., Nicholls, J., & Whitehouse, H [7].

II. DESCRIPTION OF THE SAMPLE

The focus of the study is on teachers involved in professional education. During the data collection we interviewed teachers from two member institutions of the Székesfehérvár Vocational Training Centre. The two member institutions are I.István Technikum and Vajda János Technikum. The questionnaire was sent to about 80 instructors in the two institutions and 33 of them answered the questions.

Questions were asked on two main topics, artificial intelligence and climate change. In this paper, only the questions related to climate change are examined. Responses to the questions were anonymous and voluntary. Data was collected through an online questionnaire with a deadline for completion in the third quarter of 2023.

24% of respondents were male and 76% female. The age of the respondents was asked and this information was used to categorise the teachers into generations. Based on the age data provided, the average age of respondents was 48 years. The most frequent age, i.e. the mode value of the data, was 50 years, while the median value was 49 years. From the gender and age data obtained, it is possible to see the phenomenon that is typical of the Hungarian teaching profession, namely the higher proportion of women in the profession and the low number of young, early career teachers.

In terms of generations, the majority of respondents, 67%, are from Generation X. The second largest group is Baby Boomers with 15%. Generation Y is represented with 12%, while Generation Z is represented with 6%. It can be

clearly seen that more than 4/5 of respondents are in their 50s or beyond. Knowing this, in the future, we should not only think about the succession of teachers, but also about the presence of a large generational gap between teachers and students.

In terms of place of residence, 58% of respondents live in a town, while the remaining 42% live in a village. For all generations, a higher proportion of respondents live in towns than in villages. The exception to this is Generation Z respondents, all of whom indicated a village as their place of residence.

We also looked at the marital status of respondents, which shows that 66% are married and a further 9% are in a civil partnership. 6% are single, 9% are widowed and 10% are single parents. It was also important to ask about the number of children when examining demographic variables. Having children can be a variable and a measure of thinking about the future, confidence and faith in the future. A total of 61 children were born to the 33 respondents. The responses received paint a very interesting picture. Those in Generation X had the most children, more than 80% of all children. The baby boomer generation had over 16% more children. Sadly, only slightly more than 3% of Generation Y had children. And Generation Z educators were childless. It is important to note that both Generation Y and Generation Z are still within the average age of parenthood, which in our country is currently 29 years for the first child and 30 years for the second [8]. We also looked at the number of children born per respondent in a given generation, which is exactly 2 for baby boomers and 2.23 for Generation X. For generation Y, it is only 0.5 children.

The survey also included a question on the subject area in which the responding teacher teaches. Three categories were set up for this question. Public knowledge in the real field accounted for 12% of the responses and public knowledge in the human field for 42%. The remaining 46% of the respondents came from the professional field.

III. TEACHERS' INSIGHTS ON CLIMATE CHANGE

In one block of questions, we tried to find out what the teachers who filled in the questionnaire think and what their attitude is towards climate change. One of the questions in this block was whether they think climate change is causing changes in their own lives. On this question, over 81% of respondents thought they would experience noticeable changes. Looking deeper into this question, we find that only Generation X respondents answered in the negative. It is known that members of the younger generations Y and Z are more sensitive to climate change, but all members of the oldest baby boomer generation in the study believe they will experience these impacts in their lifetime.

We asked whether they felt uncertain about climate change. Here the majority of respondents, nearly 94%, said they were uncertain about the issue. Respondents who did not feel uncertain were baby boomers and members of Generation X. The very high percentage of uncertainty makes it clear that even those with at least a tertiary education in their home country who are educated feel that they have a number of questions about climate change.

We also looked at whether teachers felt fear or anxiety when thinking about climate change. Over 63% of respondents felt fear and anxiety. The responses show that all members of Generation Z feel fear or anxiety among those who completed the questionnaire. Some of the other generations do not feel anxiety. Half of Generation Y and more than 40% of Generation X indicated a negative response.

Following the questions on uncertainty and fear, we also looked at how well informed respondents feel about climate change. Responses were given on a scale of one to ten, with 1 being not at all informed. The mean of the responses received was 6.06, while both the mode and median of the data was 7. It is an interesting correlation that the responding teachers report quite high levels of self-awareness, yet there is a high level of uncertainty in the group. Despite or because of their awareness, the feeling of uncertainty is also very high.

As before, respondents were asked to rate their preparedness for a climate change disaster on a scale of one to ten. The question did not specify what kind of disaster situation was meant, specifically to allow respondents to think in general terms. In this question, there is already more uncertainty than in the previous case. The simple arithmetic mean value of the responses is only 4.09. While the most frequently occurring item is 5. The median value only takes a value of 4.

The survey also asked respondents whether they think artificial intelligence or climate change will cause major changes in their lives and society in the next 5-10 years. When we relate the response to this question to the subject area taught, we find that all those who taught in the field of general studies in the real sciences identified climate change as having a greater impact. By contrast, a higher proportion of teachers in the humanities (64%) and vocational (60%) subjects cited artificial intelligence. Typically, generations X (59%) and Y (75%) perceive the impact of AI to be greater in their own lives, while baby boomers (60%) and generation Z (100%) perceive the impact of climate change to be stronger. Generation Z respondents all feel the impact of climate change will have a greater impact on their lives in the next 5-10 years. Interestingly, different answers were obtained when we examined the same question in a different sample of Hungarian language and foreign language teachers [9]. In that sample, all generations identified the impact of artificial intelligence as the strongest. In the current sample, all but Generation Y rated the impact of climate change as more important.

When asked which phenomenon will have the greatest impact on the economy and society as a whole, artificial intelligence clearly dominates. 50% of those who taught in the field of general studies in the real sciences mentioned AI, while 79% of those who taught in the field of general studies in the humanities thought so. 53% of respondents teaching in a vocational field felt the same. Similar to what has been discussed above, this issue has also been examined in another sample [9]. In this respect, what can be observed is that the proportion of those who consider climate change as a stronger factor has increased slightly for baby boomers (40%) and Generation X (41%) compared to teachers of Hungarian and foreign languages. In contrast, for

generations Y (25%) and Z (0%), there was a clear decrease in the proportion of those who consider climate change to have a greater impact on the economy and society.

In this respect, the answers of Generation Z to the previous two questions in the present sample deserve special attention, because they are extreme. While climate change is always seen as a stronger influence on their own lives, artificial intelligence is also always seen as a stronger influence on the economy and society as a whole.

In order to find out the views, attitudes and perceptions of VET teachers on climate change, we asked two questions, to which they were given the opportunity to give a free-word answer. In the first of these questions, we asked respondents to state what security risks climate change will pose in the next 5-10 years. For this question, 28 evaluable answers were received. In the course of the analysis, the text responses received were read through and grouped according to the focus areas indicated. It naturally follows that more than one security risk could be identified in a single response. Among the security risks, respondents identified a number of factors, but only those that appeared in at least 10% of the responses are included here. The emergence of wars and internal conflicts (10.71%) and the emergence of additional epidemics and health risks (10.71%) were identified as risks by respondents. More than 14% of responses identified the risk of imbalances in the natural environment. This fear seems well-founded and may have already been overcome if we look in more detail at the book by Ferenc Jordán [10].

Respondents consider rising temperatures to be a significant safety risk, which appears in nearly 18% of responses. As well as the emergence and increase in occurrence of weather anomalies, which occurred in over 21% of responses. The most frequently mentioned responses were immigration and migration (28.5%), food security (32.14%) and changes in water conditions and water scarcity (42.8%).

It is clear from the above data that the first two security risks are the risks of not meeting basic needs. In third place is migration and immigration, while fears of the effects of climate change are very common.

In the second free-response question, respondents were asked to briefly describe the societal risks and changes that climate change will pose in the next 5-10 years. We followed a similar methodology as for the previous question to examine the responses. Again, only responses with a value above 10% are presented in detail. A total of 27 assessable responses were received for this question.

The social risks also include epidemics and health problems (14.8%). Which, as we have seen in the case of the silica crisis, could be an unintended effect of globalisation. The risk of constraints in meeting basic needs also appears. Both water scarcity (14.8%) and food shortages (25.9%) are already significant factors in the responses. Second among the social risks identified is the increase in social conflict (33.33%), where respondents identified processes that represent internal tensions in society. Frequently mentioned ideas included "growing social differences" or "fragmenting societies, greater layers of extreme poverty". Migration was the most highly identified social risk (48.1%). It can be seen that almost half of the respondents identified migration as a source of economic - social risk.

Comparing the responses with the data from the paper on negative factors affecting the security of Hungary [5], it can be seen that when the items are taken together, the values exceed the values mentioned in the article. This increase is perhaps most noticeable in the case of migration, as in Alex Ete's paper uncontrolled migration reaches between 35 and 40%, while in our own study it exceeds 48%.

IV. Questions on the link between climate change and education

This chapter details the questions in which we asked vocational educators about their views on climate change education and whether they would be involved in such education.

We started the content part of the study by asking about concepts related to climate change, in terms of whether they appear in the background documents related to the subjects taught. These concepts included Climate Change, Sustainable Development, Green Economy, Waste Management, Climate Adaptation, Climate Resilience, etc. 69.6% of respondents indicated at least one concept that is included in educational background documents. There were negative responses from all three educational fields indicated, i.e. both the real, humanities and vocational fields indicated a negative response. The concepts concerned were also examined from the perspective that whether or not the concept is included in the educational background documents is irrelevant to whether or not it is covered in the classroom. For this question, nearly 97% indicated that at least one concept is covered in the classroom. Among the concepts covered, sustainable development stands out, with more than 75% of respondents indicating it as a concept covered.

The questionnaire also asked the responding teachers whether they thought it was necessary to prepare students for the effects of climate change. All respondents answered yes to this question. In other words, even those educators who feel less informed in the field or who think AI will have a more significant impact on the economy and society as a whole, feel that the topic is important.

We also asked how climate change knowledge would be taught, as a stand-alone subject, or in several subjects, but harmonising the knowledge elements, or whether such subjects are not needed, but the knowledge currently taught in the context of theme weeks and project days is sufficient. More than three quarters of respondents (78.79%) said that it should be presented within the framework of several subjects, coordinated and approached from different perspectives. Only 9% said that such a subject was not necessary and 12% would choose the option of a stand-alone subject. If we start to look at the responses received along generational lines, we see that only Generation X teachers said that they did not need such a subject, and that the knowledge provided in theme weeks and project days was sufficient.

The final element of the questions on climate change education was to assess the extent to which teachers in professional training would be open to teaching this type of

subject. According to the question, if there were no compulsory qualification, would you be willing to teach a subject on the phenomena and likely consequences of climate change? 57.5% of respondents to the questionnaire would be willing to teach such a subject. The generational analysis shows that half of Generation X teachers would undertake such a course, while the figure is higher for other generations (baby boomers 60%; Y 75%). 100% of Generation Z teachers would undertake such a course, confirming that this age group is more sensitive to these issues. When looking at the study by subject group, teachers of vocational subjects are the most willing to undertake such a subject (67%). Teachers of general studies in the real sciences and humanities are equally likely to undertake such a subject (half to half).

V. Conclusion

In this research, we tried to show the perceptions of participants in vocational education and training in Hungary on climate change and climate change education through a small sample. In several questions, we found that different generations of teachers have different views on climate change and its education.

However, it is also clear that the trainers believe that there is a need for education and knowledge transfer in this area. Although opinions are already divided on how to transfer knowledge.

Several international studies have been carried out on a similar topic. For example, a paper co-authored by colleagues at the Riga University of Technology, which examines, among other things, how to integrate climate change education into university education, not only in professional education, but also in a multidisciplinary way [11] [12]. But most of these papers focus on higher education, and in particular on the transfer of knowledge to students. However, if we want to reach a wider audience and provide them with the knowledge they need to tackle climate change, we need to do this at secondary level as well.

And to incorporate this knowledge into secondary education, we need to know the views of the teachers who work there. The authors of this paper consider it important to do so and to carry out further research along similar lines.

References

[1] Q. Haris Uddin and S. Syed Muzzamil Hussain, "Warming Trends over the Major Urban Centers of Pakistan Due to Climate Change," 2023 Global Conference on Wireless and Optical Technologies (GCWOT), Malaga, Spain, 2023, pp. 1-5, doi: 10.1109/GCWOT57803.2023.10064649.

[2] C. Pfister and H. Wanner, Climate and Soceity in Europe: The Last Thoussand Years Osiri Kiado 2024

[3] P. Schneider and S. J. Hook, "Global trends in lake tempe-ratures observed from space," 2012 IEEE International Geoscience and Remote Sensing Symposium, Munich, Germany, 2012, pp. 5258-5261,

[3] P. Schneider and S. J. Hook, "Global trends in lake tempe-ratures observed from space," 2012 IEEE International Geoscience and Remote Sensing Symposium, Munich, Germany, 2012, pp. 5258-5261,

[4] J. Módné Takács and M. Pogatsnik “A systematic review of Human Aspects in Industry 4.0 and 5.0: Cyber-security Awareness and Soft Skills” In: Szakál, Anikó IEEE 27th International Conference on Intelligent Engineering Systems 2023 (INES 2023) Budapest, Magyarország : IEEE Hungary Section (2023) pp. 33-39. , 7 p.

[5] A. Ete, “A magyar biztonságpercepció elemzése In Nemzet és Biztonság 2020/2. szám | 80–93. http://real.mtak.hu/116662/1/nemzet_es_biztonsag_2020_02-6_Etl.pdf (2023.10.05)

[6] Oversby, J. (2015): Teachers’ Learning about Climate Change Education; Procedia - Social and Behavioral Sci-ences, Vol. 167, P23-27, URL: https://www.sciencedirect.com/science/article/pii/S1877042814067846, letöltés ideje: 2023.11.05.

[7] Stevenson, R. B., Nicholls, J., & Whitehouse, H. (2017): “What is climate change education?”in. Curriculum Perspectives, Vol.37, p67-71, URL: https://link.springer.com/article/10.1007/s41297-017-0015-9, letöltés ideje: 2023.11.05

[8] https://www.ksh.hu/stadat_files/nep/hu/nep0007.html – (letöltés dátuma: 2024. 09. 18.)

[9] Z. Trenka, V. Varsányiné Szeredi and M. Pogátsnik “The impact and potential of artificial intelligence in language education”in MAGYAR NYELVŐR 148 : Special Issue in English (2024-5) pp. 682-695. , 14 p. (2024)

[10] Jordán F. Az ember vége, a természet esélye Open Books 2023

[11] Patlins, J. Caiko, N. Kunicina, A. Zhiravetska and V. Riashchenko, "Climate Education: Challenges of Climate Change and Energy Policies," 2020 IEEE 61th International Scientific Conference on Power and Electrical Engineering of Riga Technical University (RTUCON), Riga, Latvia, 2020, pp. 1-7, doi: 10.1109/RTUCON51174.2020.9316561

[12] Gy. Györök and B. Beszédes. “Fault-tolerant Software Solutions in Microcontroller Based Systems”. In: Orosz, Gábor Tamás (szerk.) AIS 2017 - 12th International Symposium on Applied Informatics and Related Areas organized in the frame of Hungarian Science Festival 2017 by Óbuda University : Proceedings. Székesfehérvár, Magyarország : Óbudai Egyetem (2017) 204 p. pp. 7-12. , 6 p.

The awareness of and openness to AI among vocational trainers

Viktória Varsányiné Szeredi
Doctoral School on Safety and Security Sciences
Obuda University
Budapest, Hungary
szeredi.viktoria@phd.uni-obuda.hu
ORCID: 0009-0003-9398-3009

Monika Pogatsnik
Alba Regia Technical Faculty
Obuda University
Szekesfehervar, Hungary
pogatsnik.monika@amk.uni-obuda.hu
ORCID: 0000-0002-2698-7291

***Abstract*—The existence of artificial intelligence is somewhat divisive in our society, with a negative perception for many, even in the absence of knowledge and application.**
In the age of the digital revolution, the intensity of progress varies from generation to generation, with some people using mobile phones, others using information and communication technologies on a daily basis and others identifying technological change with chat gpt. Artificial intelligence still lives as science fiction in the minds of many people, a part of our society is open to learning about it, sees it as a future possibility, but many see it as a threat that takes away our jobs and poses a security risk to humanity. As educators, the generational divide is strongly felt. This is one of the reasons why, in addition to their role as educators, they are now expected to be developmental, open-minded and forward-looking. This research investigates the awareness and openness of 33 senior teachers in two logistics-oriented technical schools towards the emergence of artificial intelligence.

Keywords—artificial intelligence, teachers attitude, education

I. INTRODUCTION

There is no doubt that the accelerated pace of development in the digital society has made continuous personal renewal essential. Industrial revolutions have brought technological innovation that has been a challenge to adapt to in every age. In the age of artificial intelligence, the emergence of automation and robots has created fear and anxiety among a section of society, mainly because many people feel that machines are a threat to their jobs [1].

As technology innovates and integrates, work processes are also changing. Artificial intelligence brings with it changes in work patterns and job losses, a realisation that has led the European Parliament to create a comprehensive set of rules to address the risks of AI through a law on artificial intelligence. It highlights the key role of education and training in enabling individuals to keep up with and adapt to technological change [2]. Developing and adapting competences and skills to expectations is also essential. This is illustrated by the introduction of online education during the COVID epidemic, which clearly revealed society's shortcomings in digital competences, both from the teacher and student perspective.

In addition to the convenience features of the AI-based applications identified, moral, ethical dilemmas and data security risks have been raised. In today's world, there is an increasing focus on privacy, so it is not only cybersecurity risks, but also the role of liability and the possibility of alienation that need to be addressed [3]. Despite the challenges, the need to meet the expectations of the times has resulted in AI being a fundamental driver, pushing the direction of development and creating a number of opportunities in education [4]. A duality can be observed, as some research has shown that skills acquired through innovative pedagogy and soft skills development, focusing on the problem, increase cybersecurity awareness [5].

The relationship between artificial intelligence and education has been examined from many perspectives, and its potential to transform personality and humanity from an ontological perspective has also been raised [6]. As a consequence, researchers are interested in a deeper examination of AI from a technology-based perspective and in terms of application models.

From an educational point of view, the possibility of personalised and needs-based learning by analysing learners' data and behavioural patterns is a positive development. Learning processes can be optimised, education can be effectively tailored to the needs of students, and there is a diversity of resources and greater accessibility [7], [8].

The analysis of the impact on the teacher-student relationship and the process of transformation is also a topical issue. With the focus on personalised learning, the role of the teacher is changing, with the creation of curricular content, assessments and feedback being coordinated by artificial intelligence. The individual teacher is no longer the sole source of learning. The change will fundamentally reorganise human relationships, it will also mean a transformation in emotional terms, and as a consequence the authority of the teacher may be reduced [9].

The loss of exclusivity in the role of knowledge and information provider for teachers, although no longer new, is frustrating and for many is difficult to deal with. Teachers are also wary of new technologies and modern applications, so in many cases they focus on why they are not good and what potential for misuse they present. Pedagogically, however, the right thing to do is to promote the successful and diverse use of digital technologies and artificial intelligence [10].

A significant number of researchers have sought to investigate and demonstrate to users the potential applications of AI in education, and to explore the current integration of AI in the digital transition. Researchers in several countries have conducted studies with a large sample

of educators to investigate their attitudes and willingness to use AI. A survey of teachers in Spain showed that although the measurement resulted in a high level of willingness to use AI, personal experience was the opposite, i.e. low [11].

A study in Serbia concluded that teachers are typically motivated by the goal of saving time as their role is changing to that of a trainer One interesting finding from a survey in the United States is that 51% of teachers have used Chat GPT [12]. Our own previous survey of language teachers, on the other hand, found 54%, although respondents indicated familiarity, which does not clearly indicate its use [13].

In our previous research, we investigated the extent to which concepts related to artificial intelligence and digital technologies, the fourth and fifth stages of the industrial revolution, appear in the content of the required and proposed curricula in vocational education [14]. Referring again to the 2019 publication of the Ministry of Innovation and Technology's VET 4.0 - A medium-term vocational policy strategy for the renewal of vocational education and training and adult learning, VET system response to the challenges of the fourth industrial revolution. Amongst other things, it stated that Industry 4.0 will result in a reduced human workforce, but new working conditions will require higher added value and the development of a knowledge-based society, and thus new and different competences will be required, which will necessitate a transformation of the VET system [15]. Unfortunately, the survey of the reformed system showed a very low coverage of search terms, which is why the idea of a new study was conceived to measure trainers' knowledge of the concepts and their attitude, perception, acceptance and orientation towards AI.

Hypothesis: Generational differences among vocational trainers significantly impact their awareness and openness toward AI, with younger generations demonstrating higher levels of openness and willingness to engage with AI-based applications compared to older generations.

II. Presentation of the Sample

In our research, the opinions and attitudes were measured of technician trainers in a slightly different way than before. Of the nearly 80 teachers of the I. István Technikum in Székesfehérvár and the Vajda János Technikum in Bicske, 33 responded to the questionnaires voluntarily and anonymously.

In the questionnaire, which was distributed online in autumn 2023 to the teachers of the schools, questions were formulated in relation to two topics of major concern to 21st century society, such as digitalization and artificial intelligence, and climate change. The focus of this research is on the perception and attitudes of educators towards digitalization and artificial intelligence.

The vast majority of the sample, 76%, is female, with the remaining 24% male. Based on the age of the respondents, they were divided into generations, with Baby Boomers represented by 15%, Generation X by 67%, Generation Y by 12% and Generation Z by 6%. The average age was 48 years, with a positive difference of 1 year between the mode and the median, which is worth comparing with the fact that the present teachers are typically teaching Generation Z, and to a lesser extent Generation ALFA, of technician students.

The distribution of respondents by place of residence did not show any significant difference, 58% live in a town, while the remaining 42% live in a village.

A high proportion of respondents, 76% in total, are married or in a civil partnership, the remainder are single, widowed or single parents. The 33 respondents in the sample have a total of 61 children, of which Generation Z has none and Generation Y has only 3%.

III. Openness, Attitude

In the first phase of the research, questions were asked to assess the marketability of the knowledge of the teachers at the two technical schools with logistics specialisations. To understand their attitudes towards digitalisation and the rise of artificial intelligence, it is relevant to map their awareness. The first question of the questionnaire asked whether they had encountered the six terms listed, which are related to 21st century technologies and education, and which are:

- Industry 4.0
- Industry 5.0
- Vocational Education and Training 4.0,
- Logistics 4.0,
- Artificial Intelligence based applications,
- Digital logistics techniques

Industry 4.0 is the fourth phase of the industrial revolution, currently underway, and is characterised by a high level of automation, robotisation and artificial intelligence. The fifth phase, already foreseen, is Industry 5.0, with a focus on sustainability and a stronger emphasis on harnessing human creativity.

Vocational Training 4.0 is a strategy to develop a marketable vocational training system to meet the needs of the industrial environment, and Logistics 4.0 is a similar fourth phase in the evolution of the logistics process and supply chain, seeking to harness AI-based applications.

12% of respondents, 4 people, were not familiar with either of these concepts. Of the sample surveyed, members of Generation X all ticked all the terms, it is worth noting that 3 out of the 4 were professional trainers. The term with the highest percentage was VET 4.0, known to 75% of the trainers, which, despite its high value, is not a positive sign, since the strategic objectives and vision of the VET system reformed a few years ago are set out in the basic document. Industry 4.0 as a term was known to almost half of the sample. The least familiar terms to the trainers were Industry 5.0 and Logistics 4.0, which are typically found in specialised research and professional journals.

67% of those who completed the questionnaire were aware of AI-based applications, which would be considered a surprising statistic in a post-Covid digital education environment and with the increasing use of chat GPT today.

The second question asked to what extent the three AI-related technology concepts of IOT, Big Data and Cloud-based applications were known or heard of by educators. Based on the answers to the previous question, despite a 67% awareness of AI, all respondents marked Cloud-based applications, i.e. 100% were aware of this technology, while the concept of Big Data was found in 50% and IOT had the lowest hit, with 12% of instructors aware of it or having heard of it.

The following questions asked how knowledgeable and open the trainers felt about digital applications and artificial intelligence, on a scale of 1 to 10.

The average score for the sample surveyed is 5.2, i.e. medium for the level of awareness of trainers on the subject. This result shows minimal variation across the three categories by subject taught, with an average of 4.8 for humanities and 5.5 for vocational subjects.

Looking at the same question by generation, the average score for the Baby Boomer generation is lower than average at 4.8 and Generation X at 5.0. Generation Z has the highest level of awareness at 6.5, and Generation Y also deviates minimally from the average with a score of 5.5. With a median of 5 and a mode of 3, the majority of respondents rated their awareness as a 3 on a scale of 10, which was also the minimum value indicated.

When examining the openness of trainers towards AI applications, a positive difference was observed, with a mean of 6.9 and a median of 7. This result symbolizes the innovative attitude of the majority of trainers, and it is also worth mentioning that 12% of the respondents, 4 trainers, indicated the maximum value, while when examining the level of awareness, none of them considered themselves well informed on the subject.

These answers can be compared to the questions in the questionnaire that asked whether the trainers had thought about and would participate in a training on AI-based applications.

Thus, in the first question, whether the instructor had thought about it, whether he/she felt the need to train in this direction, the respondents were almost half in favour, with a small positive difference, of yes. The data varied more by generation, with only 20% of Baby Boomers and a much higher proportion of Generation X, almost 60%, having considered training in this area, while Generation X and Y trainers were exactly 50/50. When this is broken down by field of teaching, vocational teachers led the way with 60% of those completing the questionnaire, with half of those in the real field at 50%, and the lowest 43% of humanities teachers considering attending training on AI-based applications.

Looking at these metrics, a very forward-looking and positive attitude is reflected in the response to the next question, which asked if they would attend training on the effective use of AI if they had the opportunity.88% of the trainers surveyed said yes. In terms of distribution, the results are also quite innovative and positive, with 20% of Baby Boomers and 14% of Generation X respondents rejecting the possibility of training only in terms of generations. If we focus on the field of education, however, the sample includes a low percentage of teachers in all three fields who would not take part in training The biggest change is for teachers with a humanities orientation, as almost 86% are open to training, as is the sample as a whole.

As a complement to our previous survey, our survey also looked at teachers' perceptions of the extent to which concepts and topics related to digitalisation and security are reflected in the background documents related to the subjects and training they teach. Of those who completed the questionnaire, 85% of respondents said that these topics were included in the documents, 54% on digital applications, 54% on data protection, 52% on sustainability and 15% on cybersecurity. We find it interesting that 1 respondent, 3% of respondents, said that none of these terms were included, and 2 respondents, 6%, did not know that these digitalisation-related terms were included. In the next question, we asked whether trainers cover these topics in their lessons, regardless of whether they are required by the training curricula. The sample examined shows a difference compared to the previous question, as all respondents cover these topics in their lessons Almost all trainers, 94% of respondents, cover digital applications in their lessons, 70% cover sustainability, half of the trainers, 51%, cover data protection and 20% cover cybersecurity.

IV. Perception of safety and vulnerability

In the next six questions of the questionnaire, the focuse was on the topic of safety, looking at the impact of the emergence of AI on educators' perceptions of the threat it poses and the impact it has on their lives.

The first question asked whether teachers in general perceive the rise of AI as a positive or negative development. Overall, the results show an open-minded attitude among teachers, with 67% of the sample surveyed perceiving it as a positive, innovative new opportunity, 12% as a negative, risk, and one fifth of respondents, 21%, could not specify. Looking at respondents by marital status, the most hesitant group was married people, In terms of generation, Generation Y rated it as an entirely positive option, while Baby Boomers rated it as a new option at 75%. Generation Z members, although only 2 respondents overall, were the least likely to see it as a major risk, while the other respondent indicated that they did not know. Members of Generation X were the most uncertain, with 27% not able to judge, 64% identifying it as a new option and the remaining 9% seeing it as a negative. In terms of gender, the proportions are very similar, with slightly more women seeing it as a positive.

The second question asked whether AI poses a security risk in general. Compared to the previous question, despite the higher proportion of positive attitudes, 85% of the sample surveyed considered it to be a security risk, 6% of respondents considered the opposite and 9% were undecided.

In the third question, we listed some options from which respondents had to choose what they perceive as the most security risk or threat for the future with the rise of artificial intelligence. The majority of respondents, 67%, see the risk in terms of the security of information and personal data, 21%

see the risk in terms of job loss and 6% see the risk in terms of the widening generation gap. The remaining 3-3% of the sample surveyed see themselves as unable or unwilling to keep up with this change, or see none of these as security risks. The results show no major differences in the sample in terms of marital status, gender or generation.

Similarly to the previous question, but in terms of work and society, 30% of the sample surveyed do not feel that their environment or education is at risk. The largest proportion of respondents, 42%, see the emergence of AI as a threat to the future generation and 21% to education.

The fourth question asked how much change artificial intelligence, the digital shift, and thus the new way of thinking, brings or requires from respondents on a scale of 1 to 10. No positive or negative aspects of this were highlighted.

The mean for the sample studied is 6.5, with a mode and median of 3. The result indicates that they perceive it to be bringing them more change than the average.

The last two questions of the questionnaire were designed to measure the respondents' perceptions on the AI side of the two major topics, but the last two questions were targeted to ask in which category the trainers felt more at risk.

Firstly, we looked at how teachers perceive climate change or AI to make a bigger difference in their own lives in the next 5-10 years. The sample showed little variation, with 54% of respondents feeling the impact of artificial intelligence and 46% feeling the impact of climate change as a factor that will turn things around for them. In terms of the distribution by subject taught, unlike the sample as a whole, those teaching in the real sciences see climate change as a major change.

In terms of the economy-society unit, the spread of artificial intelligence was again identified by respondents as a change with a greater impact, this time even higher at 64%. Focusing on the professional field, the vast majority of humanities teachers, 79%, identified AI as the major impact, while real and professional teachers were evenly split. In terms of generations, 100%, 75% of Generation Y, and 60% of Generation X and Baby Boomers feel it is a more powerful variable. In terms of gender, the women's group showed the largest difference, with both perceiving the impact of AI on themselves and society as a whole as the most influential.

V. SUMMARY

Our study builds on previous research that examined the extent to which vocational education and training (VET) integrates the VET 4.0 strategy and the characteristics of the 21st-century digital environment into its curricula. The findings revealed that while topics such as the various stages of industrial revolutions, artificial intelligence technologies, and intelligent logistics systems are less prominent, there is a stronger focus on the security aspects of digitalization technologies.

In the current research, we extended this inquiry by exploring teachers' awareness of and attitudes towards technological change, with particular emphasis on their ability and willingness to engage with AI-based applications, as well as their perceptions of associated threats or safety.

Despite the evolving role of teachers and their continued influence as models of openness and innovation for students, there is a notable lack of research in this area within the existing literature.We consider the investigation of teachers' openness to and awareness of emerging technologies and innovations to be a pertinent area of research. Unfortunately, national regulations restricting mobile phone usage in schools do not appear to achieve the intended outcomes, largely due to inadequate access to necessary technological resources. The rapid advancement of technology presents an inherent advantage to younger generations. Rather than imposing restrictive measures, a more progressive approach may involve fostering safe usage, cultivating creativity, and preparing future generations to meet the demands of the evolving labor market.

The attitudes of the teachers are evident in the response rate, with only 40% of educators from the two technical schools choosing to participate in this brief, anonymous survey, which required 5-10 minutes of their time.

Unfortunately, the responses reinforced our hypothesis that generational differences influence the ability to adapt to technological changes, particularly with respect to safety concerns. While the teachers who completed the survey generally demonstrate openness, a willingness to evaluate changes, and a positive outlook, their knowledge on the subject requires further enhancement.

While the study provides valuable insights into vocational trainers' awareness and attitudes toward AI, its findings are constrained by the small sample size of 33 respondents. This limited sample may affect the generalizability of the results, suggesting that further research with a larger, more diverse group of participants is necessary to confirm and expand upon these observations

REFERENCES

[1] S. Sinha, B. Kaur and Sheenam, "Impact of Artificial Intelligence on Work," 2023 7th International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech), Kolkata, India, 2023, pp. 1-6, doi: 10.1109/IEMENTech60402.2023.10423502

[2] European Parliament, A mesterséges intelligencia használata és veszélyei, https://www.europarl.europa.eu/topics/hu/article/20200918STO87404/a-mesterseges-intelligencia-hasznalata-es-veszelyei

[3] P. Xiaoling, "Discussion on Ethical Dilemma Caused by Artificial Intelligence and Countermeasures,"2021 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC), Dalian, China, 2021, pp. 453-457, doi: 10.1109/IPEC51340.2021.9421271.

[4] Y. Zheng, H. Meng and W. Jia, "Application Research and Challenges of Artificial Intelligence in Primary and Secondary Education,"2022 12th International Conference on Information Technology in Medicine and Education (ITME), Xiamen, China, 2022, pp. 162-166, doi: 10.1109/ITME56794.2022.00043.

[5] J. Módné Takács , M. Pogátsnik The Presence of Cybersecurity Competencies in the Engineering Education of Generation Z, ACTA POLYTECHNICA HUNGARICA 2024, 21 : 6 pp. pp.107-127.

[6] X. Li, Z. Ma, Y. Tu and Y. Du, "Study on the Application of Artificial Intelligence Technology in Empowering Education : Taking "Intelligent Learning Partner" as an Example," 2021 2nd International Conference on Information Science and Education (ICISE-IE), Chongqing, China, 2021, pp. 1449-1453.

[7] S. Kuzenkov, "The Possibilities of Using Artificial Intelligence Technologies in Education and Science," 2024 4th International Conference on Technology Enhanced Learning in Higher Education (TELE), Lipetsk, Russian Federation, 2024, pp. 88-91.

[8] M. Pan, J. Wang and J. Wang, "Application of Artificial Intelligence in Education: Opportunities, Challenges, and Suggestions," 2023 13th International Conference on Information Technology in Medicine and Education (ITME), Wuyishan, China, 2023, pp. 623-627.

[9] F. Kamalov, D.Santandreu Calonge, I. Gurrib New Era of Artificial Intelligence in Education: Towards a Sustainable Multifaceted Revolution. Sustainability. 2023; pp. 15 12451.

[10] Z. Szűts, Z. Námesztovszki, "A digitalizáció kihívásai a civil, mindennapi felhasználó szemszögéből," CIVIL SZEMLE, 2023, vol 5, pp. 57–66.

[11] H.Galindo-Domínguez, Héctor, N. Delgado, L. Campo. D. Losada, (2024). Relationship between teachers' digital competence and attitudes towards artificial intelligence in education. International Journal of Educational Research. pp. 126. 102381.

[12] V. Kuleto, MP. Ilić, R. Bucea-Manea-Ţoniş, D-F. Ciocodeică, H. Mihălcescu, V.Mindrescu The Attitudes of K–12 Schools' Teachers in Serbia towards the Potential of Artificial Intelligence. Sustainability. 2022; pp. 14, 8636.

[13] Z. Trenka, V. Varsányiné Szeredi., M. Pogátsnik The impact and potential of artificial intelligence in language education Magyar Nyelvőr 148. 2024:pp. 682–695

[14] Z. Trenka. V. Varsányiné Szeredi, M. Pogátsnik A szakképzés-fejlesztés és szakmai pedagógusképzés összefüggései : XI. Trefort Ágoston Szakképzés- és Felsőoktatás-pedagógiai Konferencia. Tanulmánykötet: Konferencia helye, ideje: Budapest, Magyarország 2023.11.09. Budapest: Óbudai Egyetem, 2023, pp 33-48

[15] Innovációs és Technológiai Minisztérium: Szakképzés 4.0 – A szakképzés és felnőttképzés megújításának középtávú szakmapolitikai stratégiája, a szakképzési rendszer válasza a negyedik ipari forradalom kihívásaira, https://ikk.hu/files/Szakkepzes_4.0_VI.pdf, 2019, pp. 7.