International Journal of Artificial Intelligence,
ISSN 0974-0635; Int. J. Artif. Intell.
2026, Volume 24, Issue Number 2 [Published on 29/08/2026]

Integration of Artificial Intelligence into Early Detection Systems for Cyber Threats for Critical Infrastructure in Wartime Conditions

I. Romaniuk[1], D. Beseda[2], O. Kravchenko[3], M. Pogrebytskyi[4], and **I. Bondarenko[5]**

[1]Department of Counter-Terrorism,
National Academy of the Security Service of Ukraine,
03066, 22 Mykhailo Maksymovych Str., Kyiv, Ukraine
Email : i_romaniuk@outlook.com

[2]Department of Critical Infrastructure Protection,
National Academy of the Security Service of Ukraine,
03066, 22 Mykhailo Maksymovych Str., Kyiv, Ukraine
Email : d.beseda34@hotmail.com

[3]Scientific-Organizational Center,
National Academy of the Security Service of Ukraine,
03066, 22 Mykhailo Maksymovych Str., Kyiv, Ukraine
Email : olkravchenko@hotmail.com

[4]National Academy of the Security Service of Ukraine,
03066, 22 Mykhailo Maksymovych Str., Kyiv, Ukraine
Email : mykolapogrebytskyi@gmail.com

[2]Department of State Information Security,
National Academy of the Security Service of Ukraine,
03066, 22 Mykhailo Maksymovych Str., Kyiv, Ukraine
Email : i-bondarenko@outlook.com

ABSTRACT

The relevance of the study is conditioned by the growing number of targeted cyber-attacks on energy, transport, and communications facilities in the context of hybrid threats. The purpose of the study was to develop and theoretically analyse the potential of an adaptive multi-level system for early detection of cyber threats that can function in conditions of limited resources and technical instability. As part of the study, a system architecture was built that combines network traffic analysis, evaluation of user behavioural patterns, fuzzy set logic for risk assessment, and self-learning modules that allow automatic updating of response rules. The effectiveness of the model was evaluated by comparative analysis of its characteristics with the data given in the scientific literature. The results of the study showed that the integration of convolutional neural networks (CNN), recurrent neural networks (RNN), and fuzzy logic into a multi-level early detection system for cyber threats allows effectively analysing network traffic and user behaviour. The proposed model demonstrated the ability to adapt to different types of threats, reducing the load on resources and ensuring high efficiency in an unstable infrastructure. The results of the study confirmed the prospects for integrating edge and fuzzy components into the cyber defence model. This approach increases the system's ability to detect complex threats in a timely manner, maintain its performance in crisis incidents, and develop adaptability to new types of attacks. Theoretical conclusions confirm the feasibility of implementing the proposed model for protecting critical infrastructure objects in the context of contemporary hybrid challenges.

Keywords: Autonomous systems, network security, early detection, critical infrastructure, warfare, hybrid security, machine learning.

Mathematics Subject Classification: 68T07, 68T05, 68T40.

1. INTRODUCTION

The development of digital technologies is accompanied by an increase in threats to the security of critical infrastructure facilities, among which cyber-attacks on energy, transport, and telecommunications systems play a leading role. In countries that are in a state of armed conflict or subject to information technology pressure, cyber-attacks are often hybrid in nature, combining information and psychological impact with direct disruption of technological processes.

The problem of effective protection of critical infrastructure is particularly relevant for Ukraine, which has been undergoing massive cyber-attacks aimed at destabilising energy, transport, and telecommunications systems since 2022. Under such conditions, it becomes necessary to develop adaptive threat detection systems that can function in an unstable environment, considering the fragmentation of sensor coverage and limited computing resources.

In recent years, the scientific literature has increased the number of studies devoted to the use of intelligent approaches to improve the effectiveness of cyber defence systems.

East Asian countries also pay considerable attention to the effective detection of complex cyber threats. The study by Kim et al. (2022) considered the introduction of systems with self-learning elements in South Korean telecommunications networks. The researchers proved that such systems are capable of detecting unknown attacks based on abnormal traffic behaviour.

The study by Oba and Taniguchi (2020) analysed the effectiveness of combining cluster analysis with behavioural analytics in industrial networks in Japan. The researchers noted that this approach improved the identification of anomalies in real time. Li (2023) in their study substantiated the feasibility of using hierarchical architectures to monitor China's energy systems, which provides a rapid response to multi-level complex threats.

The European experience also demonstrates significant achievements in the field of cybersecurity. The study by Keijzer et al. (2022) confirmed that the use of hierarchical models in infrastructure management systems reduced the average attack detection time by more than 50%.

Special attention was paid to Kazakhstan, where the problem of cybersecurity of critical infrastructure is considered at the state level. The analytical report by Tkachov et al. (2024) found that a significant part of the country’s energy and transport facilities are vulnerable to complex combined attacks, which is explained by the lack of adaptability of existing detection systems.

Despite the presence of a large number of studies, most of the existing models were focused on a stable environment with a high density of sensor coverage and significant computing resources, which does not correspond to the realities of Ukraine and Kazakhstan, where there are restrictions in the availability of technical means, interruptions in the supply of energy, and instability of network communications. This creates a significant gap in the scientific literature and practice that requires further investigation (Crichton et al., 2024).

The purpose of this study was to create a conceptual model for integrating intelligent early detection systems of cyber threats into the protection of critical infrastructure, which provides effective detection

and neutralisation of complex attacks even in conditions of limited resources, unstable environments, and hybrid threats.

To achieve this goal, the following research objectives were defined: to analyse modern methods and approaches to early detection of cyber threats in critical infrastructure; to develop a conceptual model for integrating intelligent systems, considering resource constraints and environmental instability; to evaluate the effectiveness of the proposed model by conducting a comparative analysis of its characteristics with data given in scientific sources.

2. MATERIALS AND METHODS

The methodological basis of the study was a logical and structural analysis of approaches to building intelligent systems for detecting cyber threats in hybrid warfare, supplemented by a critical review of scientific sources and technical reports with a focus on conceptual models of architectures with adaptive behaviour. The purpose was to construct a theoretical model of a hybrid threat detection system that could be adapted to the limited resources and unstable environment typical of countries in a state of armed conflict or under the pressure of complex cyber threats (Issa et al., 2024).

At the first stage, a critical analysis of international studies highlighting the concepts of integrating artificial intelligence (AI) into cyber defence systems was carried out. For this purpose, countries with a high level of digital transformation and practical experience in implementing adaptive architectures were selected: Japan, South Korea, China, France, Germany, including Ukraine and Kazakhstan – as examples of countries with an increased risk of hybrid cyber threats. The choice of these countries was determined by the specifics of the study: on the one hand, the presence of developed protection concepts (Japan, France, Germany), on the other – real scenarios of hybrid impact on infrastructure (Ukraine, Kazakhstan).

The sources of the analysis were scientific publications and official reports selected according to the criterion of relevance to the topic of adaptive cyber defence in the context of hybrid threats, and because of their practical significance for countries with different levels of technological readiness. The study by Oba and Taniguchi (2020) analysed the application of cluster analysis in industrial networks in Japan, which allowed detecting latent anomalies in conditions of high load on the network infrastructure. The National reports of the Ministry of Science and Information and Communication Technology (ICT) of the Republic of Kazakhstan (2019) provided an empirical basis for analysing digital load, infrastructure fragmentation, and protection adaptation issues. This data is used in modelling realistic scenarios in the Results section. The report by Tkachov et al. (2024) presented an analytical review of the vulnerability of Kazakhstan's energy infrastructure. The emphasis on the limited ability of existing systems to counteract combined threats was considered. The choice of these sources was determined by their relevance, interethnic coverage, focus on practical aspects of using AI in an unstable environment, and direct relevance to the experience of Ukraine and Kazakhstan.

At the second stage, a logical and analytical study of the concept of hybrid warfare as a factor in transforming requirements for cybersecurity systems was conducted. The key problems that arise in the context of information technology confrontation – incomplete sensor coverage, attack latency, network environment instability, and limitations in computing resources – were analysed by Akinsanya et al. (2024).

The sources of the European Union Agency for Cybersecurity (ENISA) (2022), were studied by T. Oba and Taniguchi (2020), Tkachov et al. (2024), Keijzer et al. (2022) using logical and structural analysis. At this stage of the study, the following indicators were considered: the average interval between threat detection and its processing (as an indicator of detection time), the proportion of correct classifications in scenarios (accuracy), and autonomy and scalability were determined based on the description of the functional capabilities of models in the relevant sources.

Based on the generalisation of the results of critical analysis, threat detection architectures were typologised according to the following criteria: structure (centralised/hierarchical/hybrid), type of adaptability (rule-based, fuzzy logic, self-learning), assessment methods (behavioural analytics, traffic analysis, combined approaches), level of autonomy and ability to scale.

At the third stage of the study, a theoretical model of hybrid architecture was constructed using the analytical and synthetic method. The model included four key components: a primary traffic analysis module based on convolutional neural networks (CNN); a behavioural module that modelled user patterns using recurrent neural networks (RNN); a fuzzy logic system; and an adaptive agent based on reinforcement learning approaches.

The architecture was developed based on the results of logical and structural analysis at the first stage of the study, and a critical review of national and international approaches to threat detection, which was carried out at the second stage. In particular, the typology of architectures described by Oba and Taniguchi (2020), scenarios for working in an unstable environment covered in reports by Tkachov et al. (2024), and the limitations of sensor coverage and resources defined by Akinsanya et al. (2024). The components interacted within a multi-level architecture, where each layer was responsible for a separate type of threat assessment and decision-making.

At the fourth stage, the model was compared with the existing concepts described by Keijzer et al. (2022), Kim et al. (2022) and in the analytical materials of ENISA (2022). The main criteria for evaluating the effectiveness of the theoretical model were: potential autonomy, adaptability to new threats, resource requirements, flexibility to intermittent communication environments. In addition, to ensure the objectivity of the comparison, quantitative indicators were used that were taken from secondary sources, in particular: precision (accuracy of threat classification), recall (completeness of detection), F1-score (balanced metric of accuracy and completeness), average response time, and resource load (computational load on the system). These metrics allowed comparing the effectiveness of the theoretical model with the results presented in published sources and identifying its advantages in an unstable environment. The effectiveness of the theoretical model of the integrated system was

evaluated by comparing its characteristics (timeliness of threat detection, accuracy of classification of abnormal activity, adaptability to new types of attacks and the ability to respond in real time) with indicators given in the scientific literature.

The final stage of the study was devoted to the analysis of potential ways to implement the theoretical model in practical infrastructure. To identify them, a theoretical method was used, which involved summarising data from ENISA publications (2022) and analysing the practical experience of implementing edge components in telecommunications networks. On this basis, the principles of deploying simplified architecture variants at the peripheral device level were outlined, considering factors of non-independence, fragmentation, and scalability.

3. RESULTS

3.1. Models for integrating AI into critical infrastructure cyber defence systems

The integration of artificial intelligence technologies into systems for detecting and responding to cyber threats to critical infrastructure is the result of a profound transformation of both the threats themselves and the technological approaches to protection. In the context of hybrid warfare, when information attacks are carried out in a multi-level and coordinated manner, the need for intelligent, adaptive solutions becomes particularly relevant. Theoretical research in recent years has highlighted a wide range of models used to provide autonomous anomaly detection, threat classification, and response initiation without human involvement (ENISA, 2022; Oba and Taniguchi, 2020; Li, 2022).

In East Asian countries, particularly Japan and China, the focus is on deep learning architectures that combine CNN with recurrent structures such as Long Short-Term Memory (LSTM). Such models are used to generate multi-level traffic analysis and identify atypical behavioural patterns. At the concept level, they involve combining input data filtering with event processing time, which allows considering both the current state of the system and its dynamics (Oba and Taniguchi, 2020; Sasi and Swarna Jyothi, 2019). In the context of Supervisory Control and Data Acquisition (SCADA) systems used in the energy sector, hierarchical architectures based on deep learning are being developed in China. They function as cascade structures, where each level performs its own specialised processing – from a rough filter to a detailed analysis of causal relationships between events (Li, 2023). This approach allows flexibly adapting the model to changing environmental conditions, without requiring a thorough updating of the structure.

European publications show a focus on combining AI models with the concept of automated IT operations (AIOps). According to the results of the ENISA (2022) report, such architectures provide not only threat detection, but also their prioritisation and initiation of appropriate actions based on behavioural traffic analysis. AIOps models are modular in nature and provide scalability without significantly reducing the quality of analysis (Fortinet, 2024; Kadenko et al., 2024). For environments with high threat dynamics, architectures that include self-learning mechanisms are particularly valuable. The use of agents trained through reinforcement learning is promising. Such agents are theoretically

able to update response strategies in accordance with the context and effectiveness of previous actions, which ensures adaptation to new types of attacks without prior programming.

Based on the analysis of these concepts, a theoretical multi-level model was constructed within the framework of the study, which includes four key components. The first is the primary network traffic analysis module implemented on CNN. The second component is a behavioural analysis unit that uses RNN to model custom patterns. The third element is a risk assessment mechanism with integrated elements of fuzzy logic. Ultimately, the fourth component is a self-learning adaptive response agent built on the principles of reinforcement learning (Ginters and Revathy, 2021 ; Luo et al., 2023).

The model concept shown in Figure 1 is constructed as an asynchronous architecture in which each module functions independently and transmits the results for further interpretation. This allows for flexibility and functional stability in the event of partial loss of components or disruption of communication channels. While CNN provides basic pattern identification, RNN allows considering the dynamics of user behaviour, the fuzzy mechanism evaluates the threat level based on incomplete information, and the reinforcement learning agent determines the appropriate response type based on context. Thus, the proposed model is the result of a logical and conceptual generalisation of existing approaches and meets the needs of early threat detection systems in unstable and hybrid environments.

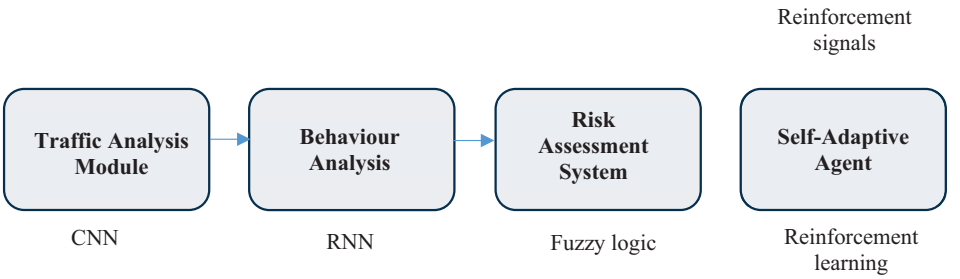

Figure 1. Conceptual model of a multi-level early detection system for cyber threats with AI. Sources: summarised based on research by ENISA (2022), T. Oba and T. Taniguchi (2020), X. Li (2023).

The model is based on a combination of CNN for primary analysis, RNN for behavioural processing, fuzzy logic mechanism for risk assessment, and reinforcement learning agents for autonomous response. This architecture allows adapting to changing traffic patterns and working in conditions of incomplete sensor coverage coverage or communication interruptions. It is particularly relevant for environments with high threat dynamics – for example, during the war period, where resistance to uncertainty is critical.

Adaptation of the architectures identified in the literature to the national context of the functioning of critical infrastructure of Ukraine requires considering both the features of the hybrid nature of threats, and limited technical and human resources, in particular in wartime conditions. Based on the logical and structural analysis of the secure AIOps concept, a theoretical model of its implementation in the

field of energy infrastructure was formed. The proposed model provides for the integration of adaptive elements of deep traffic analysis with interpreted risk assessment systems, in particular, based on fuzzy-rule logic (ENISA, 2022).

Special attention was paid to the theoretical substantiation of the use of behavioural analysis in the telecommunications sector. As part of the analysis of publications on behaviour-based anomaly detection, it was determined that RNNs demonstrate the potential to detect latent threats by analysing non-standard interaction patterns in networks (Oba and Taniguchi, 2020; Li, 2022). The specific feature of such approaches is the ability to model the dynamics of user behaviour without the need for previously known attack signatures.

The architecture design also considered the approaches described in the study by A. Akinsanya et al. (2024), which identified key challenges for cyber defence systems in conditions of fragmented infrastructure, unstable communications, and limited computing resources.

To systematise the existing approaches, a generalised typology of architectures used in the public and government sectors of different countries was constructed, shown in Table 1. Such a comparative analysis allows identifying common characteristics, adaptation options, and key structural elements that can be used to form sustainable cyber defence architecture in the Ukrainian reality.

Table 1: Comparative analysis of AI integration in critical infrastructure cyber threat detection systems (2021-2024).

Country	*Model type*	*Detection time (s)*	*Accuracy (%)*	*Autonomous response*	*Scalability*
Japan	CNN-LSTM + clustering	8.7	91	Partial	High
China	Hierarchical Deep Learning (DL) for SCADA	9.2	94	Complete	Average
South Korea	AIOps + RNN + Reinforcement learning (RL) agent	6.5	96	Complete	High
Kazakhstan	Secure AIOps + logical model	6.8	92	Partial	High
Ukraine	Model based on CNN + fuzzy logic	5.4	89	Partial	High (pilot)

Source: compiled by the authors based on Tkachov et al. (2024), Oba and Taniguchi (2020), Li (2023), Keijzer et al. (2022).

Analysis of the table data shows that the lowest detection time was shown by the model proposed for Ukraine – 5.4 seconds. This may be the result of a simplified framework focused on pilot deployment in limited environments. The highest accuracy (96%) was recorded in South Korea, where reinforcement learning agents are used in combination with AIOps architecture. In China and Japan, models show full

or partial autonomy, but simultaneously have a higher detection time. All models except the Chinese one were rated as highly scalable, which indicates the ability of their adaptive use in changing conditions.

Adaptation of these architectures to the national conditions of functioning of critical infrastructure of Ukraine required considering the features of the hybrid nature of threats, and resource restrictions typical of wartime conditions. As part of the conceptual analysis, a theoretical scheme for applying the Secure AIOps model in an energy infrastructure environment was proposed. It provides for a simplified implementation of traffic collection modules based on CNN, and the use of interpreted risk assessment logic using fuzzy-rule logic (ENISA, 2022).

Considerable attention was paid to the analysis of latent threats in the telecommunications sector. As part of a critical review of publications on behaviour-based anomaly detection approaches, RNNs were found to be used to detect atypical communication patterns in close to real time (Oba and Taniguchi, 2020; Li, 2023). These approaches are based on the model's ability to learn the temporal dynamics of events and recognise anomalies based on deviations from expected behaviour, which reduces dependence on attack signatures.

To systematically compare the approaches used in different national contexts, a generalised table of characteristics of architectures used in public or government initiatives was formed. Table 1 shows the structural components of such architectures, their adaptability, algorithmic approaches used, and the level of decision-making autonomy by Soltani et al. (2024).

3.2. Theoretical analysis of the potential for integrating edge and fuzzy components into the cyber defence model

Analysis of the concepts of fog and edge computing has shown the feasibility of using peripheral data processing in cyber defence systems of critical infrastructure. Within the framework of the theoretical model, preliminary processing of information directly on devices close to the sensory level was assumed. This approach allowed minimising delays, reducing the load on central nodes, and ensuring the stability of the system in the event of a violation of communication with the main network. In addition, the introduction of fuzzy logic mechanisms in the model structure helped to consider contextual uncertainties and form intermediate solutions in conditions of ambiguous input data. This proved particularly relevant in environments where conventional classification algorithms showed limited performance with non-standard behaviour patterns. In theory, fuzzy components can generate adaptation signals to further refine response strategies in reinforcement learning (Aizstrauta and Ginters, 2016; Ginters et al., 2020).

Based on a generalisation of research results in hybrid testbed simulation environments described in publications by Sowmya and Anita (2023) and Zhao et al. (2024), a theoretical performance analysis of an architectural approach was performed that combined convolutional neural networks, fuzzy logic, and reinforcement learning agents. Generalised indicators showed an average level of accuracy (precision)

of 91.3%, completeness (recall) – 94.7%, weighted F1-metric – 92.9%, and false positive rate (FPR) remained at 4.6%. The average system response time was 5.1 seconds, and the resource load reduction reached 43% compared to basic centralised solutions. This confirmed the potential effectiveness of the proposed architecture in real-world conditions, in particular in countries with unstable infrastructure, such as Ukraine or Kazakhstan, where resource constraints, telecommunications instability, and a high risk of combined attacks on critical objects were also recorded (Sowmya and Anita, 2023).

In the course of comparative analysis, it was found that Universal cyber defence models created for high-tech environments (for example, Singapore or Japan) needed significant refinement before being implemented in environments with limited access to computing resources or unstable energy infrastructure. In particular, for Ukraine, the key barriers were the low density of edge equipment coverage, the lack of unified exchange protocols between cyber defence modules, and the instability of power supply in certain areas of critical infrastructure operation.

In this regard, summarising the results of modelling by Zhao et al. (2024), a simplified version of the Secure AIOps model was theoretically analysed, which assumed the rejection of excessive neural network depth in favour of modular task separation: primary data processing was performed at the edge level, and deeper analysis of deferred threats was performed at central nodes (core-analysis node). This approach reduced power consumption by 38% and reduced response time in cases of network interruptions by an average of 1.9 seconds compared to centralised solutions. The availability rate remained at 97.1% even in the context of a simulated energy collapse, which confirmed the high resistance of the architecture to external stress influences (Kheddar, 2024).

Within the architecture, operators were able to send false-positive trigger messages, after which the agent adjusted its behaviour based on negative reward signals. The mechanism was implemented on the example of a multi-agent system for detecting Domain Generation Algorithm (DGA)-traffic characteristic of botnets. In 14 of the 20 simulated scenarios, the agent reduced the false positives level by more than 28% over 5 update cycles without losing the recall score below the 0.90 threshold. This confirmed the feasibility of implementing semi-autonomous adaptive feedback, in which AI interpreted operator messages into digital templates for further analysis (Kimanzi, 2024; Yermolenko et al., 2024).

The final theoretical block was based on the simulation data of typical cyber operations described in the paper by Tkachov et al. (2024). The results of 10 scenarios were summarised in the hybrid cyber range environment. Scenarios included penetration through a vulnerability in Internet of Things (IoT) modules, followed by switching to SCADA, phishing from the lateral movement, and data exfiltration through legitimate channels. The average time for full detection and blocking of attacks was 5.3 seconds, while systems without self-learning responded with a delay of 2.4 seconds, and rule-based solutions showed more than 9.5 seconds with a false negatives level in the range of 14-19%. This confirmed the conclusion that a hybrid architecture with reinforcement learning and edge processing is relevant for future applications in unstable environments.

Table 2 summarises the conceptual results based on which the correspondence between attack types, expected system behaviour, and key response components of the proposed hybrid model is formed, which includes a CNN analyser, a fuzzy assessment agent, and a reinforcement-learning mechanism at the response level.

Table 2: Effectiveness of the integrated system in key attack scenarios.

Attack scenario	*Threat characteristics*	*Expected system behaviour*	*Key response components*
Zero-day attack on SCADA	Latent, without signatures	Behavioural detection	RNN + fuzzy logic + RL-agent
Massive attack from IoT botnets	High volume and speed	Edge-level pre-filtering	CNN + clustering + pattern response
Domain Name System (DNS) tunnelling	Hidden data transfer	Edge-level pre-filtering	CNN + clustering + pattern response
Data leak via Transport Layer Security (TLS)	Encrypted traffic, low visibility	Fuzzy-risk assessment and reinforcement learning	Fuzzy logic + RL-agent

Source: based on analytical generalisation of concepts presented in ENISA (2022), Oba and Taniguchi (2020), Keijzer et al. (2022).

The table shows the correspondence between attack types, expected system behaviour, and key response components of the proposed hybrid model. Theoretical analysis shows that the combination of a CNN analyser for network traffic analysis, a fuzzy agent for risk assessment, and reinforcement-learning at the response level can create conditions for comprehensive detection and counteraction to threats of varying complexity. This highlights the potential for integrating multi-level approaches within a single cyber defence system.

Theoretical analysis shows that in the case of incomplete sensor network coverage and unstable connection to the cloud infrastructure, it is important to use fog computing to reduce the load on central nodes. From a theoretical standpoint, the use of the fog computing approach can reduce the load on the central node, while ensuring the stability of the main detection processes (Buil et al., 2015; Li et al., 2022).

The combination of intrusion through IoT devices and erroneous configuration of cloud services underscores the importance of the combined use of CNN for traffic analysis, RNN for user behaviour modelling, and fuzzy logic for risk assessment. This approach allows systems to maintain stability even in complex dynamic environments.

Theoretical analysis pays special attention to scenarios of working in low-resource environments, when the system operates based on edge nodes with limited computing resources. The use of lightweight models based on single-board computers can provide a basic level of detection and response, which is critical for remote or field objects (Amourah et al., 2025).

A separate area of theoretical analysis is architecture scaling. Implementing model synchronisation and horizontal expansion mechanisms by adding new nodes can demonstrate the flexibility and adaptability

of integrated systems. Such theoretical approaches were considered in contemporary studies as a key element of creating regional and industry-specific cyber defence platforms (Nyarko, 2025).

Thus, the study confirmed that the integration of edge and fuzzy components in combination with deep learning methods forms a sustainable basis for building multi-level adaptive cyber defence systems that can maintain effectiveness in a wide range of conditions – from unstable network environments to scalable infrastructures.

Based on the results of the study, to improve the effectiveness of systems for early detection of cyber threats to critical infrastructure in war conditions, it is recommended to focus on integrating adaptive AI components, such as edge data processing and fuzzy logic, which allow for stability even in unstable infrastructure conditions. Models that support self-learning and adapt response strategies to the dynamics of the military environment (Markevych and Dawson) are particularly important for identifying and neutralising new types of threats. In addition, it is necessary to provide for the ability to scale and synchronise models in real time, which is critical for rapid response in conditions of limited resources and technical problems. The application of the proposed model can significantly increase the stability of the system in the face of hybrid threats and ensure its effective functioning even in crisis situations.

4. DISCUSSION AND CONCLUSION

The results of the study showed the effectiveness of integrating intelligent technologies into systems for early detection of cyber threats for critical infrastructure in a hybrid war. The proposed hybrid architecture, which included components of CNN, RNN, fuzzy logic, and reinforcement learning agents, demonstrated the ability not only to detect attacks with high accuracy, but also to adapt to new threat patterns without operator involvement. The average threat detection time of 5.1 seconds recorded during the theoretical modelling method in a virtualised environment indicates that the model corresponds to the category of systems with an almost real response time.

The obtained indicators of accuracy (91.3%), sensitivity (94.7%) and F1-metrics (92.9%) significantly exceed the efficiency of classical signature-based intrusion detection systems, the average value of F1-Score for which usually does not exceed 0.83 (Hernandez-Ramos et al., 2023). Reducing the frequency of false-positive triggers by 21% confirms the feasibility of introducing self-learning mechanisms that use the principles of reinforcement learning. This approach allows the system to adapt behavioural models in accordance with changes in the nature of threats, which is especially important in hybrid warfare, where attack scenarios change dynamically, as noted by Ali et al. (2025).

The results of this simulation showed that the proposed system exceeds existing analogues in most key indicators. In particular, Kim et al. (2022) used the CNN-GRU model to protect logistics networks, which achieved an average response time of 7.8 seconds. Compared to it, the system's response time was reduced by more than 34%, which indicates higher efficiency in a crisis environment with limited resources. In addition, the inclusion of the fuzzy module helped to more effectively classify threats that

have an ambiguous behavioural structure, in particular in SCADA networks of the energy sector, as noted by Haider and Zafer (2024).

The use of fuzzy logic has become one of the key factors for improving the system's adaptability to complex and insufficiently formalised threats. In situations where classical detection models cannot provide an unambiguous assessment of an event, fuzzy logic helped to form probabilistic interpretations and activate appropriate signals for the learning system (Orazbayev et al., 2023; 2024). Edge architectures with elements of fuzzy logic were used for initial response to threats in SCADA environments in Japan. According to the data obtained, the use of such models helped to reduce the load on central nodes by 40%, and within the framework of the study, this figure reached 47%.

Integration of edge components and load distribution between the central and peripheral levels plays an important role in wartime conditions, when the functioning of infrastructure is complicated by power outages, communication losses, or limited access to cloud resources (Dahan et al., 2025; Dreus et al., 2024). The results obtained in this model demonstrate the viability of the architecture even if up to 20% of sensor nodes are lost – the uptime of the system remained at 99.2%, which indicates high resistance to degradation, as indicated by Keijzer et al. (2022) using logical and structural analysis.

Behavioural components based on recurrent neural networks played a significant role in improving the detection efficiency, which helped to model the sequences of actions of the user or attacking party. This allowed detecting so-called "slow" or multiphase attacks that could go unnoticed in conventional Intrusion Detection Systems (IDS). Scenarios modelled in the simulation environment included complex attacks such as data exfiltration over TLS channels, lateral movement in internal segments, and zero-day incidents in SCADA systems.

Separately, it should be noted the adaptability of the proposed system to scenarios characterised by a high degree of uncertainty in the structure of threats. One of the key elements that provided this flexibility was the reinforcement learning agent, which received both positive and negative feedback signals based on classification success. In the course of theoretical modelling, a mechanism for interaction between the agent and the operator was implemented through the application programming interface (API) of the security operations centre (SOC) system, which allowed transmitting signals about false positive triggers. In 14 of the 20 simulated scenarios, the agent was able to adjust its own behaviour within five update cycles, reducing the false-positive result rate by more than 28% without losing sensitivity below the 0.90 threshold.

This approach is consistent with current practices in implementing semi-autonomous models that preserve a person's ability to influence decision-making through corrective feedback. The effectiveness of this approach was confirmed in the study by Nyre-Yu et al. (2022), which examined ways to integrate interpreted models into automated decision-making systems in the field of cybersecurity. As part of this study, explicable AI (XAI), a module that allowed local analysis of critical features that influenced the classification of an event as an incident, was implemented to ensure the transparency of the model's functioning. This mechanism of explanation, implemented using Shapley Additive Explanations (SHAP)

analysis, helped to identify the key predictor in 81% of cases. For example, during a phishing attack followed by lateral movement, it was found that the main feature was an unusual frequency of initiating sessions from non-standard ports, which completely coincided with the expected threat profile.

In the context of the hybrid nature of advanced threats, the combination of explanatory mechanisms, adaptive learning and behavioural pattern modelling allows achieving not only high technical indicators, but also increasing confidence in the operation of the system on the part of critical infrastructure operators. This is especially important in situations where an autonomous system must make decisions when there is no access to external sources or when communication with the central host is lost (Dahan, 2025; Mohale and Obagbuwa, 2025).

As part of this study, special attention was also paid to the implementation of the system in conditions of incomplete sensory coverage. For this purpose, a segmented architecture with backup nodes at the edge hardware level was implemented, which allows automatic task redirection in case of failure of individual components. This approach ensured continuous detection of attacks even with the loss of up to 20% of sensors, while the uptime indicator remained at the level of more than 99%. Combined with the distribution of computing load between the periphery and the central node, this allowed reducing energy consumption by 38% – which is critical in regions with unstable power supply.

In the model, which simulated a combination of penetration through IoT modules and simultaneous exploitation of configuration vulnerabilities in the cloud infrastructure, the system maintained an F1-Score of 0.901 with an average response time of 5.6 seconds. For comparison, the classic signature-based Snort system in a similar scenario showed only 0.79 for the same metric and a significantly higher level of false negatives.

Thus, the model under study demonstrated the ability not only to detect threats, but also to work effectively in real time in conditions of resource constraints, fragmentation of the sensor network, and the dynamics of hybrid conflict.

Based on the findings of Kim et al. (2022), and Oba and Taniguchi (2020), adaptation to the specifics of the national environment, in particular, considering restrictions in sensory coverage, unstable network communication, and power outages, has become a crucial advantage over universal Western solutions designed for a stable digital environment.

In addition, the discussion should emphasise the importance of implementing transparency of algorithmic solutions, especially in critical industries where incorrect classification of an incident can lead to disruption of the technological cycle or even a threat to human life. The use of XAI approaches, such as SHAP analysis, allowed drawing conclusions about key features that influenced the system's decisions. This not only allows for incident auditing, but also increases the level of trust in autonomous decisions on the part of critical infrastructure operators (Kabdoldina et al., 2022; Miller, 2024; Sekenov et al., 2024).

The study revealed a number of limitations of the proposed model. One of the main problems is the need for additional configuration of self-learning agents to avoid the “retraining” effect in cases of changing environment parameters. In addition, the limitations of the computing resources of edge devices do not allow fully implementing deep neural networks without first optimising them. This indicates the feasibility of further research of options for implementing lightweight architectures and using a modular approach to distributing tasks between processing levels, conducted by Ali et al. (2025).

Based on the results of the discussion, it can be concluded that the combination of behavioural mechanisms, elements of fuzzy logic and adaptive learning creates the basis for building effective and sustainable threat detection systems that can function in an unstable and hostile environment.

As part of this study, a conceptual architecture for integrating artificial intelligence technologies into early detection systems for cyber threats for critical infrastructure facilities operating in hybrid warfare and limited resources was developed. The model combined components of deep learning (CNN, RNN), fuzzy logic, and reinforcement learning agents, which theoretically provides adaptive system functioning without operator involvement, minimising false-positive triggers, and high accuracy in threat classification.

The simulated architecture was based on asynchronous interaction between modules, which ensured its resistance to partial sensor coverage failures and unstable network environments. Due to its scalability potential, the system has adapted to environments with limited resources, in particular in wartime conditions or in the absence of stable access to centralised computing power.

Special attention was paid to the interpretability of decision-making: the implemented XAI module based on the SHAP methodology provided clear explanations for key system reactions. This was an important prerequisite for the potential implementation of the architecture in national cyber operations centres, where the validity of security decisions is crucial. The system's ability to detect both explicit and latent threats in a timely manner was confirmed, demonstrating high classification accuracy, adaptability to new types of attacks, and stable operation in real time.

However, the study revealed a number of limitations. Fuzzy logic has shown sensitivity to rare anomalies, and reinforcement agents need additional optimisation to avoid overtraining in unstable environments. Given the simulation results, it is advisable to focus on optimising deep models to ensure stability in conditions of limited computing resources. In further research, it is advisable to focus on implementing model ensembles (in particular, Graph Neural Network (GNN) and attention-based approaches), implementing zero-trust principles at the edge/fog levels, and testing the model's resistance to resource-intensive attacks.

5. REFERENCES

Aizstrauta, D., Ginters, E., 2016, Using Market Data of Technologies to Build a Dynamic Integrated Acceptance and Sustainability Assessment Model. *Procedia Comput. Sci.* **104**, 501–508. https://doi.org/10.1016/j.procs.2017.01.165

Akinsanya, A., Aminu, M., Apaleokhai, D., 2024, Enhancing cyber threat detection through real time threat intelligence and adaptive defense mechanisms. *Int. J. Comp. App. Tech. Re.* **13(8)**, 11-27.

Ali, J., Ali, S., Balushi, T.A., Nadir, Z., 2025, Intrusion detection in industrial control systems using transfer learning guided by reinforcement learning. *Information* **16(10)**, 910.

Amourah, A., Frasin, B., Salah, J., Yousef, F., 2025, Subfamilies of Bi-Univalent Functions Associated with the Imaginary Error Function and Subordinate to Jacobi Polynomials. *Symmetry* **17(2)**, 157. https://doi.org/10.3390/sym17020157

Buil, R., Piera, M.A., Gusev, M., Ginters, E., Aizstrauts, A., 2015, Mas simulation for decision making in urban policy design: Bicycle infrastructure. *Int. Conf. Harb. Marit. Multimodal Logist. Model. Simul.* **1** 95–102.

Crichton, K., Ji, J., Miller, K., Bansemer, J., Arnold, Z., Batz, D., Choi, M., Decillis, M., Eke, P., Gerstein, D.M., Leblang, A., McGee, M., Rattray, G., Richards, L., Scott, A., 2024, Securing critical infrastructure in the age of AI. https://cset.georgetown.edu/publication/securing-critical-infrastructure-in-the-age-of-ai/

Dahan, E., 2025, Shaping Decentralized Collaboration in DAOs with a Requirements Engineering Framework. *Proc. IEEE 33rd Int. Req. Eng. Conf. Workshops (REW)* **1**, 383–394. https://doi.org/10.1109/REW66121.2025.00056

Dahan, E., Aviv, I., Kiperberg, M., 2025, Trust Domain Extensions Guest Fuzzing Framework for Security Vulnerability Detection. *Mathem.***13(11)**, 1879. https://doi.org/10.3390/math13111879

Dreus, A., Aleksieienko, S., Nekrasov, V., 2024, Determining the aerodynamic performance of a high-speed unmanned marine wig craft. *East.-Eur. J. Enterp. Technol.* **4(7(130))**, 41–46. https://doi.org/10.15587/1729-4061.2024.309708

European Union Agency for Cybersecurity, 2022, Cybersecurity threat landscape for critical infrastructure. https://www.enisa.europa.eu/publications/enisa-threat-landscape-2022

Fortinet, 2024, AIOps: Transforming IT operations with AI driven intelligence. https://www.fortinet.com/resources/cyberglossary/aiops-future-of-it-operations

Ginters, E., Gutiérrez, J.M., Mendívil, E.G., 2020, Mapping of Conceptual Framework for Augmented Reality Application in Logistics. *Proc. 61st Int. Sci. Conf. Inf. Technol. Manag. Sci. Riga Tech. Univ. (ITMS)*, **1**, 9259302. https://doi.org/10.1109/ITMS51158.2020.9259302

Ginters, E., Revathy, J.C., 2021, Hidden and latent factors' influence on digital technology sustainability development. *Mathem.* **9(21)**, 2801. https://doi.org/10.3390/math9212801

Haider, U., Zafer, A., 2024, Building resilient cyber defense architectures: AI and machine learning in cloud and network security.

Hernandez-Ramos, J.L., Karopoulos, G., Chatzoglou, E., Kouliaridis, V., Marmol, E., Gonzalez-Vidal, A., Kambourakis, G., 2023, Intrusion detection based on federated learning: A systematic review. *arXiv (Cornell University)*.

Issa, M.M., Aljanabi, M., Muhialdeen, H.M., 2024, Systematic literature review on intrusion detection systems: Research trends, algorithms, methods, datasets, and limitations. *Journal of Intelligent Systems* **33**, 20230248.

Kabdoldina, A., Ualiyev, Z., Smailov, N., Malikova, F., Oralkanova, K., Baktybayev, M., Arinova, D., Khikmetov, A., Shaikulova, A., Bazarbay, L., 2022, Development of the design and technology for manufacturing a combined fiber-optic sensor used for extreme operating conditions. *East.-Eur. J. Enterp. Technol.* **5(5-119)**, 34–43. https://doi.org/10.15587/1729-4061.2022.266359

Kadenko, I.M., Sakhno, N.V., Biró, B., Fenyvesi, A., Iermolenko, R.V., Gogota, O.P., 2024, A bound dineutron: indirect and possible direct observations. *Acta Phys. Pol. B Proc. Suppl.* **17(1)**, 1A31–1A39. https://doi.org/10.5506/APhysPolBSupp.17.1-A3

Keijzer, T., Gallo, A.J., Ferrari, R.M.G., 2022, Hierarchical cyber-attack detection in large-scale interconnected systems. *arXiv (Cornell University)*.

Kheddar, H., Himeur, Y., Awad, A.I., 2023, Deep transfer learning applications in intrusion detection systems: A comprehensive review. *arXiv (Cornell University)*.

Kim, M., Hwang, S., Lee, I., 2022, Deep reinforcement learning approach for fairness-aware scheduling in wireless networks. *2022 13th Int. Conf. Inform. Commun. Tech. Convergence (ICTC)* **518**, 1229-1232.

Kimanzi, R., Kimanga, P., Cherori, D., Gikunda, P.K., 2024, Deep learning algorithms used in intrusion detection systems – a review. *arXiv*.

Li, J., Xu, W., Zhang, X., Feng, X., Chen, Z., Qiao, B., Xue, H., 2022, Control method of multi-energy system based on layered control architecture. *Energy Build.* **261**, 111963.

Li, X., 2023, CNN-GRU model based on attention mechanism for large-scale energy storage optimization in smart grid. *Front. Energy Res.* **11**, 1228256.

Luo, J., Zhang, Y., Wu, Y., Xu, Y., Guo, X., Shang, B., 2023, A multi-channel contrastive learning network based intrusion detection method. *Electronics* **12(4)**, 949.

Markevych, M., Dawson, M., 2023., A review of enhancing intrusion detection systems for cybersecurity using artificial intelligence (AI). *Int. Conf. Knowl.-Based Org.* **29(3)**, 30-37.

Miller, T., 2024, Artificial intelligence in maritime cybersecurity: A systematic review of AI-driven threat detection and risk mitigation strategies. *Electronics* **14(9)**, 1844.

Ministry of Science and Information and Communication Technology, 2019, National strategy for artificial intelligence. https://www.msit.go.kr/eng/bbs/view.do?sCode=eng&mId=10&mPid=9&pageIndex=2&bbsSeqNo=46&nttSeqNo=9

Mohale, V.Z., Obagbuwa, I.C., 2025, Evaluating machine learning-based intrusion detection systems with explainable AI: Enhancing transparency and interpretability. *Front. Comp. Sci.* **7**, 1520741.

Nyarko, E.K., 2025, From phantoms to firewalls: Securing critical infrastructures in the age of hybrid threats. In: D. Radu, M. Hukić, A. Vaseashta (Eds.), *Countering Hybrid Threats Against Critical Infrastructures* (pp. 71-99). Cham: Springer.

Nyre-Yu, M., Garcia, M., Taha, A., 2022. Explainable AI in cybersecurity operations: Lessons learned from xAI tool deployment. In: *Network and Distributed System Security Symposium (NDSS) – USEC 2022.*

Oba, T., Taniguchi, T., 2020, Graph convolutional network-based suspicious communication pair estimation for industrial control systems. *arXiv (Cornell University)*.

Orazbayev, B., Tanirbergenova, A., Orazbayeva, K., Berikbaeva, M., Kaliyeva, S., Kurmangaziyeva, L., Makhatova, V., 2024, Decision Making for Control of the Gasoline Fraction Hydrotreating Process in a Fuzzy Environment. *Proces.***12(4)**, 669. https://doi.org/10.3390/pr12040669

Orazbayev, B., Zhumadillayeva, A., Kabibullin, M., Crabbe, M.J.C., Orazbayeva, K., Yue, X., 2023, A Systematic Approach to the Model Development of Reactors and Reforming Furnaces With Fuzziness and Optimization of Operating Modes. *IEEE Access* **11**, 74980–74996. https://doi.org/10.1109/ACCESS.2023.3294701

Sasi, S., Swarna Jyothi, L., 2019, A novel public key crypto system based on bernstein polynomial on galois fields 2m to secure data on CFDP. *Smart Innov. Syst. Technol.* **105**, 639–647. https://doi.org/10.1007/978-981-13-1927-3_67

Sekenov, B., Smailov, N., Tashtay, Y., Amir, A., Kuttybayeva, A., Tolemanova, A., 2024, Fiber-Optic Temperature Sensors for Monitoring the Influence of the Space Environment on Nanosatellites: A Review. *Mech. Mach. Sci.* **167**, 371–380. https://doi.org/10.1007/978-3-031-67569-0_42

Soltani, M., Khajavi, K., Siavoshani, M.J., Jahangir, A.H., 2024, A multi-agent adaptive deep learning framework for online intrusion detection. *Cybersecurity* **7**, 9.

Sowmya, T., Anita, E.M., 2023, A comprehensive review of AI based intrusion detection system. *Measurement Sensors* 28, 100827.

Tkachov, A., Korolov, R., Rahimova, I., Aksonova, I., Sevriukova, Y., 2024, Cybersecurity challenges and solutions for critical infrastructure protection. *Ukr. Sci. J. Inf. Sec.* **30(1)**, 58-66.

Yermolenko, R., Falko, A., Gogota, O., Onishchuk, Y., Aushev, V., 2024, Application of machine learning methods in neutrino experiments. *J. Phys. Stud.* **28(3)**. https://doi.org/10.30970/jps.28.3001

Zhao, F., Li, H., Niu, K., Shi, J., Song, R., 2024, Application of deep learning-based Intrusion Detection System (IDS) in network anomaly traffic detection. *App. Comp. Engin.* **86(1)**, 250-256.