The object of the study is a corporate network with a dynamic structure and centralized management. The subject of the research is the processes of ensuring the protection of information resources in the corporate network. The goal is to develop a method of protecting information in the corporate network. The development is based on the Zero Trust Security strategy, according to which access to the network is allowed only after verification and identification of information. The task is to develop an effective method of protecting information resources and managing cyber security in the corporate network, taking into account the complex aspects of malicious influence. The following results were obtained. It is shown that the complex, diverse presentation of information in the network requires a comprehensive approach with the division of mixed content of information into segments according to the target orientation. Based on CISA's (Cybersecurity and Infrastructure Security Agency) Zero Trust Maturity Model, a method of targeted traffic segmentation is proposed. It allows detailed analysis of the interaction between applications, users and corporate network infrastructure, which increases the level of complex threats detection by 15 %. A method of protecting information resources of a socio-cyber-physical system is proposed, which, based on the principle of the Zero Trust Security strategy, improves the monitoring and management of cyber security of information resources by taking into account social aspects. This allows to detect and respond to threats in real time and adapt security policies according to the dynamics of user behavior and general security conditions. Integrating analytical methods and modern technologies into a security strategy creates a foundation for adaptive and resilient cyber defense

Keywords: cybersecurity, protection of information resources, security policy; semiotic model, socio-cyber-physical system, confidentiality, integrity, authenticity, intelligent analysis, traffic control

UDC 004.056
DOI: 10.15587/1729-4061.2024.313158

DEVELOPMENT OF A METHOD FOR PROTECTING INFORMATION RESOURCES IN A CORPORATE NETWORK BY SEGMENTING TRAFFIC

Maksym Tolkachov
Associate Professor
Department of information systems named after V. O. Kravets**
Nataliia Dzheniuk
Associate Professor*
Serhii Yevseiev
Corresponding author
Doctor of Technical Science, Professor, Head of Department*
E-mail: Serhii.Yevseiev@gmail.com
Yurii Lysetskyi
Doctor of Technical Science, Associate Professor, General Director
Subsidiary "SNT Ukraine"
Akademika Palladina ave., 44A, Kyiv, Ukraine, 03142
Volodymyr Shulha
Doctor of Historical Sciences, Senior Researcher, Rector
State University of Information and Communication Technologies
Solomianska str., 7, Kyiv, Ukraine, 03110
Ivan Grod
Doctor of Physical and Mathematical Sciences, Professor
Department of Mathematics and Methods of its Teaching
Ternopil Volodymyr Hnatiuk National Pedagogical University
Maksyma Kryvonosa str., 2, Ternopil, Ukraine, 40027
Serhii Faraon
PhD, Associate Professor
Department of Cyber Warfare
The National University of Defense of Ukraine
Povitrianykh Syl ave., 28, Kyiv, Ukraine, 03049
Ihor Ivanchenko
PhD, Associate professor
Department of Security Information Technology
National Aviation University
Liubomyra Huzara ave., 1, Kyiv, Ukraine, 03058
Igor Pasko
PhD, Senior Research
Scientific-Research Center of Missile Troops and Artillery
Herasyma Kondratieva str., 165, Sumy, Ukraine, 40021
Dmytro Balagura
PhD, Associate Professor
Department of Information Technology Security
Kharkiv National University of Radioelectronics
Nauky ave., 14, Kharkiv, Ukraine, 61166
*Department of Cyber Security**
**National Technical University "Kharkiv Polytechnic Institute"
Kyrpychova str., 2, Kharkiv, Ukraine, 61002

Received date 02.08.2024
Accepted date 07.10.2024
Published date 23.10.2024

How to Cite: *Tolkachov, M., Dzheniuk, N., Yevseiev, S., Lysetskyi, Y., Shulha, V., Grod, I., Faraon, S., Ivanchenko, I., Pasko, I., Balagura, D. (2024). Development of a method for protecting information resources in a corporate network by segmenting traffic. Eastern-European Journal of Enterprise Technologies, 5 (9 (131)), 63–78. https://doi.org/10.15587/1729-4061.2024.313158*

1. Introduction

Gathering information to assess the security of information systems in real time is relevant in today's world of rapidly evolving threats. Recently, new types of threats have emerged, such as socio-cybersystemic attacks, targeted cyberattacks, and insider intruders. Existing approaches to dividing information flows, for example, using static one-

step tagging in Zero Trust Security [1], are insufficient for further analysis of information system protection, which is confirmed by recorded cases of successful cyber-attacks. In addition, this approach to system protection does not take into account the peculiarities of relatively new types of information systems, such as the Internet of Things and the Industrial Internet of Things, social networks and the data they process. As a result, new approaches to ensuring the security of information systems appear, in particular, approaches based on constant dynamic monitoring of information system behavior.

Organizations have learned to collect large amounts of data related to system behavior and security, and open sources of security data collected by various organizations have emerged. In order to properly process these data, it is necessary to correctly separate them and sort them for convenient analysis by SIEM (Security information and event management) and UEBA (User and entity behavior analytics) class systems.

There are currently a number of commercial solutions in the SIEM and UEBA industry, including QRadar SIEM, ArcSight Enterprise Security Manager (ESM), Securonix UEBA, Micro Focus Security ArcSight User Behavior Analytics (UBA), and others. The practical experience of applying such solutions shows that their current implementations are not always effective in the case of existing tagging, traffic division and subsequent security monitoring in the Zero Trust Security approach [1]. This also applies to the analysis of security incidents, the identification and analysis of the causes and consequences of offensive actions and the response to security incidents. Semiotic models and techniques developed on their basis can be used to expand the possibilities for providing information to and from such systems.

In addition, new methods of intelligent data processing have appeared. Artificial intelligence (AI)-based systems have advanced significantly, and this carries a number of risks regarding its impact on the resilience of various systems' defenses. However, with proper control, AI systems can mitigate security threats, their consequences, and control them [2, 3]. Due to the complexity and presence of many influence channels, targeted attacks represent the main threat to security. Moreover, they influence all components of security services. In particular, it is confidentiality, integrity, authenticity, availability of traffic management, involvement service.

The questions of applying the intelligent data analysis methods in the tasks of adequate division of the transmitted information for security assessment [4, 5] and the question of the transmitted data correct dynamic division for the further process of network protection [6] remain open. The collected data must be presented in a convenient form to determine a complete and adequate set of metrics that reflect the security situation. When controlling the information transmitted, the social influence on the content of the traffic is not taken into account, which can have a significant impact on the security and performance of the network. There is also no analysis of the semiotic (symbolic) aspects of the traffic, which can help identify potential threats at the level of meaning and context of the data.

Among the available methods of intellectual data processing, the semiotic approach is promising from the point of view of the described problem. It makes it possible to structure existing knowledge about the subject area, to trace dependencies between various objects, processes and events. This allows to draw conclusions about the causes and consequences of various, including social, events based on the identified dependencies.

Thus, the threat of inefficient use of existing algorithms, deployment structures and new attacks require a new methodology for directly forming security systems. Multi-contour security systems [7–9] should be used with an approach based on the concept of Zero Trust Security [10] covering the entire network security stack. The developed approach should combine an improved analysis structure, dynamic change of access policies, dynamic analysis using a semiotic approach for reliable protection against the latest challenges of cyber threats in cyberspace.

2. Analysis of literary sources and problem formulation.

The authors in [11] consider the impact of human behavior on cybersecurity, proposing a new approach to measuring and evaluating this impact through the implementation of the Human Error Evaluation and Reduction (HEART) methodology in the Plan-Do-Study-Act (PDSA) cycle. The main benefits of the study are highlighting the need for attention to the human factor in cyber security and developing a framework for quantifying this impact, which expands the understanding of how behavioral aspects affect the protection of information systems. However, there are potential difficulties in applying and interpreting the HEART technique without specialized training or experience, and further research is needed to confirm the effectiveness of the proposed approach.

Work [12] focuses on establishing security standards for information systems, including recommendations for risk management, privacy protection, and access control. Key aspects include the systematization of security controls, their application methods, and procedures for assessing compliance. Risk Management Guidelines in NIST SP 800-53 Rev. 4 cover the processes of identifying, assessing, prioritizing, and responding to risks to ensure the security of federal information systems. It is important to implement a risk management framework that includes regular monitoring and evaluation of security controls and adapting risk management strategies to changes in threats, vulnerabilities and impacts on the organization. The recommendations emphasize the need to involve the organization's management in the risk management process and use risk-based decisions to make informed decisions about security [12]. Privacy protection and access control recommendations include the development and implementation of policies and procedures that ensure the protection of personal information and limit access to information resources to only those individuals who have a need for such access. Access control includes identification and authentication mechanisms, management of access attributes, and monitoring and logging of access attempts. But at the same time, issues related to the possibility of integration with social engineering methods and obtaining signs of hybridity and synergism are not taken into account. Privacy protection requires the implementation of data management policies, minimization of data collection, transparency of data use, and mechanisms for requesting access or personal data correction [12]. The systematization of security controls includes the classification of controls by functional categories, such as access control, risk assessment, system and

information integrity, physical security, identification and authentication. Control is divided into groups according to their purpose, which allows organizations to systematically implement security measures. This includes basic control functions that are mandatory for all systems and additional ones that can be selected depending on the needs of the organization and risk assessment.

The paper [13] considers the use of ultra-broadband signal technology to protect information in cyber-physical space. The advantages include high reliability and immunity of the system, effective masking of the structural signal, ability to work in conditions of high-level interference. Disadvantages are related to the high complexity of system implementation, potential limitations in compatibility with existing equipment, as well as the need for specialized knowledge to develop and maintain such systems. This approach provides effective data protection in cyber-physical systems, but requires detailed analysis and planning when designing protection systems.

In [14], an attack using weaknesses in the implementation of the Acorn RISC Machine (ARM) TrustZone technology, which is widely used to ensure the security of critical tasks in IoT (Internet of Things), is analyzed. The authors discovered that TrustZone's trustlet or OS (Operating System) download verification procedure may use the same verification key and lack proper mechanisms to prevent rollbacks. The attack and its possible consequences for device security are described in detail. The article presents the results of experiments conducted on popular devices, showing the universality of the problem. The disadvantages of the article are that the implementation of the proposed solutions can be difficult due to the complexity of the chain of trust and the distribution of devices. The conclusions are based on experiments with a limited number of devices, which may not fully reflect the state of affairs in the entire industry.

Work [15] examines in detail the concept of a double contour protection in socio-cyber-physical systems, emphasizing the importance of an integrated approach to security, which includes both internal and external protection measures. It highlights the relevance of the problem in the context of a wide range of threats, from social engineering to cyber attacks, and offers a comprehensive approach to risk assessment and the development of defensive strategies. However, the implementation of such an approach can be difficult due to the need for deep integration between various system components and require significant resources. The article contains an in-depth analysis of modern security threats and challenges, and offers an innovative approach to assessing and improving the level of security. It also highlights the importance of integration between different system components for effective protection. The disadvantages of this work are the possible complexity of implementation in real conditions due to high requirements for system integration and limited demonstration of the practical effectiveness of the proposed concept due to the lack of detailed case studies.

For a comprehensive review of the multi-contour protection topic and its application based on the concept of Zero Trust Security, it is worth pointing to the following study [16]. This article is relevant because interactivity technologies between users, devices and cloud services are moving beyond the traditional network, significantly expanding the perimeter and creating gaps in control, making organizations more vulnerable to attack. It is described how after intruders penetrate the perimeter, their further movement in the network becomes uncontrollable. The article analyzes in detail the concepts of attacks and their application in such areas as software security, mobile device security, and moving target protection (MTD). The article transfers the concept of attacks to the network level as a criterion for assessing the resistance of networks to potential zero-day attacks. Abstraction levels for software are analyzed, their compatibility with the attack surface model and resources inside the network is described. Proposed heuristic algorithms for this compatibility.

A necessary component of the protection system is the use of artificial intelligence (AI) systems, which have been significantly improved in recent years, and their capabilities have expanded. It is important to mention the National Institute of Standards and Technology (NIST) report entitled "Artificial Intelligence Risk Management Framework" (AI RMF 1.0), published in January 2023 [17]. Created jointly by the public and private sectors, the report aims to improve the ability to ensure reliability in the design, development, use and evaluation of AI products, services and systems. It emphasizes that AI technologies have a huge potential for sustainable development. But there are also potential risks associated with AI that could have negative effects on individuals and society in both the short and long term. The risks associated with AI systems are unique, so mitigating them can be a serious challenge for organizations. At the same time, if they are not solved effectively, they can lead to undesirable consequences. In [17], it is stated that the complexity of AI systems and the environment in which they are used make it difficult to identify and eliminate failures. This makes the process even more complex, as social dynamics and human behavior are often subject to change. Advances in AI can change the current state of cybersecurity and, conversely, how cybersecurity of AI systems affects their safe and secure design, deployment, and operation. Therefore, it is important to manage AI risks by following responsible AI practices. Part 2 describes four specific features that help organizations minimize the risks associated with AI systems in practice. Management, measurement and control functions are divided into different categories. Because the framework was developed through a consensus-based, open, transparent process, it reflects an integrated approach to AI-influenced systems.

The use of metrics and ontology of semiotics is well described in the article [18]. The proposed semiotic model of cyberspace allows for a more in-depth analysis of the interaction between humans and technologies, which is important for understanding and managing complex information security systems. The developed integral indicator of potential threats for the owners of network information resources provides an opportunity to improve monitoring and management of cyber security. The model presented in the article [18] divides information content into interconnected levels, which helps in assessing the level of security for each of them. Taking into account these aspects ensures the development of effective strategies for the protection of information resources in cyberspace. The developed models and methods make it possible to improve threat monitoring and cyber security management, which is critically important in the global digitalization era. Thus, taking social and perceptual aspects into account can complicate analysis and forecasting, as these aspects are subjective and may change over time. The semiotic model and integrated threat index can be

difficult to implement in practice without proper training and technical support. Despite the verification, the model needs additional research and testing in different conditions to confirm its universality and effectiveness.

3. The aim and objectives of the study

The aim of the study is to develop a method of protecting information resources in the corporate network by traffic segmentation, which consists of two stages – dynamic primary traffic macro-segmentation and micro-segmentation, which uses a semiotic approach. This approach using parameters of semiotic analysis – syntactic, semantic and pragmatic allows for flexible adaptation to changes in network activity in real time.

To achieve the goal of the work, it is necessary to solve the following tasks:

– to analyze users' behavioral data and devices for macro-segmentation of traffic by data values, their context, social aspects of traffic, relationships between traffic components and their impact on network security;

– to form a strategy of two-level dynamic marking of data to improve analysis while protecting information resources;

– to develop a method of protecting the mixed content of information of a socio-cyber-physical system based on semiotic analysis, which increases the level of information resources protection and provides flexibility in cyber security management;

– to verify the developed method of socio-cyber-physical system information mixed content protection based on semiotic analysis.

4. Research materials and methods

4. 1. Protection based on CISA's Zero Trust Maturity Model

The object of research is a network based on a dynamic, centrally controlled structure – a network with software-defined access [9, 10]. The subject of the research is the processes of ensuring the protection of information resources in the corporate network. The security of the entire infrastructure with such a network is built on the unified approach defined in CISA's Zero Trust Maturity Model.

In CISA's Zero Trust Maturity Model, network security is built on the integration of several aspects:

– continuous authentication of users;

– device status monitoring;

– data transparency through encryption and marking;

– application protection from hypervisors to containers;

– analysis of network events using artificial intelligence for personalized response to incidents;

– automation of threat response processes;

– visibility and analytics for precise access control and effective enforcement of security policies.

This integrated system provides multi-level protection that adapts to changing conditions and threats, increasing the level of infrastructure security (Fig. 1) [10].

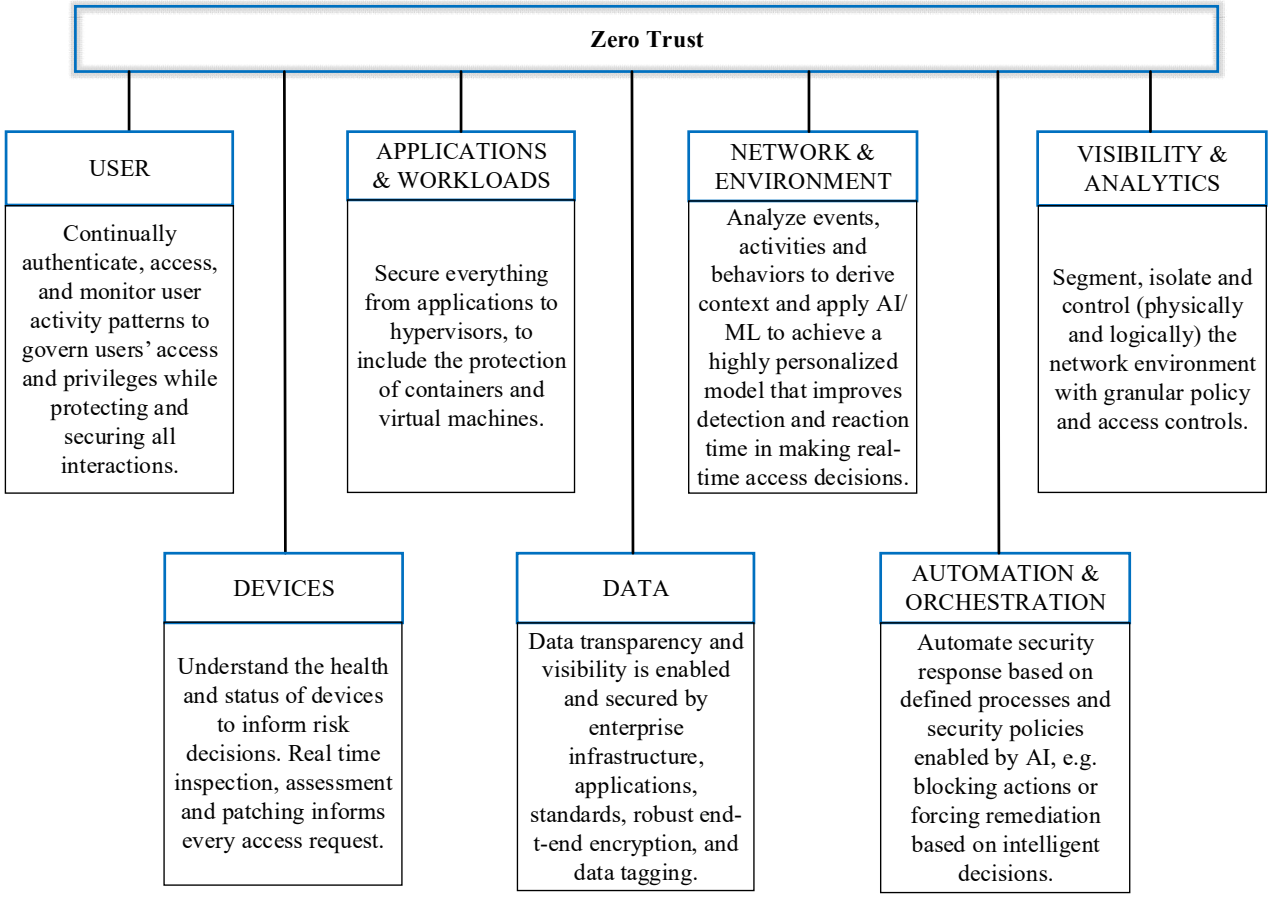

Fig. 1. Zero Trust Maturity Model

To access the network, it is necessary to be able to separate the network and traffic to simplify management and deeper dynamic analysis of information. By default, all credentials are untrusted. Comprehensive security monitoring provides insight into a breach of trust and the ability to report a breach. At the same time, the user, device, API (Application Programming Interface) and application that corresponds to this network component should be defined.

It is recommended that the network ID be used as an access policy enforcement point. Using identification technologies, a list of devices is found out, which applications generate traffic, which users and which components of information are in the load.

Changes in traffic marking begin with the macro-segmentation of the network into zones. First, a trust zone is created, and then much smaller implicit trust zones are created within the zone. First, macro segments (for example, the IoT segment) are created, in which micro-segmentation is carried out. At the same time, traffic passes between macro segments.

Based on the received information and the Zero Trust Maturity Model (ZTMM), user-to-app data is analyzed. When establishing trust, the first step is to eliminate implicit trust. Thus, when a violation is detected, it is localized by means of segmentation, marking and more specific dynamic security policies with a reduction in localization time (Table 1).

The management level organizes compliance with policies by applying the necessary plan and performs orchestration.

4. 2. Network architecture for traffic segmentation control

A network architecture based on CISA's Zero Trust Maturity Model has been developed to control traffic segmentation.

The interaction of components in this architecture follows a certain cycle that provides two-level dynamic marking of data to improve analysis while protecting information resources (Fig. 2).

The presented architecture is based on an approach that has stages of situational content at the level of access for the first stage of marking – macro-segmentation.

Marked data traffic at the edge of the network undergoes additional inspection, where under certain conditions a decision is made to change access policies and start micro-segmentation of the current session. The architecture supports requests for access to protected resources, which are performed through the control plane, where confidentiality, integrity, authenticity, traffic management availability, and involvement service are considered (Fig. 3). At this level, orchestration occurs, and dynamic security policies are applied. More detailed policies, based on the semiotic model, are applied when trust is breached at the Secure Gateway level.

Table 1

Traffic marking and segmentation

Stages	Description
Identification and classification	Each data packet is identified by type, source, destination or content
Marking-based segmentation	Traffic is divided into different segments or "channels" based on defined criteria
Prioritization and optimization	Traffic is optimized based on its priority
Security and filtering	Using segmentation and marking, potentially dangerous traffic that does not comply with established security policies is identified and blocked
Monitoring and analytics	Combining marking and segmentation allows to get a detailed overview of traffic transmission

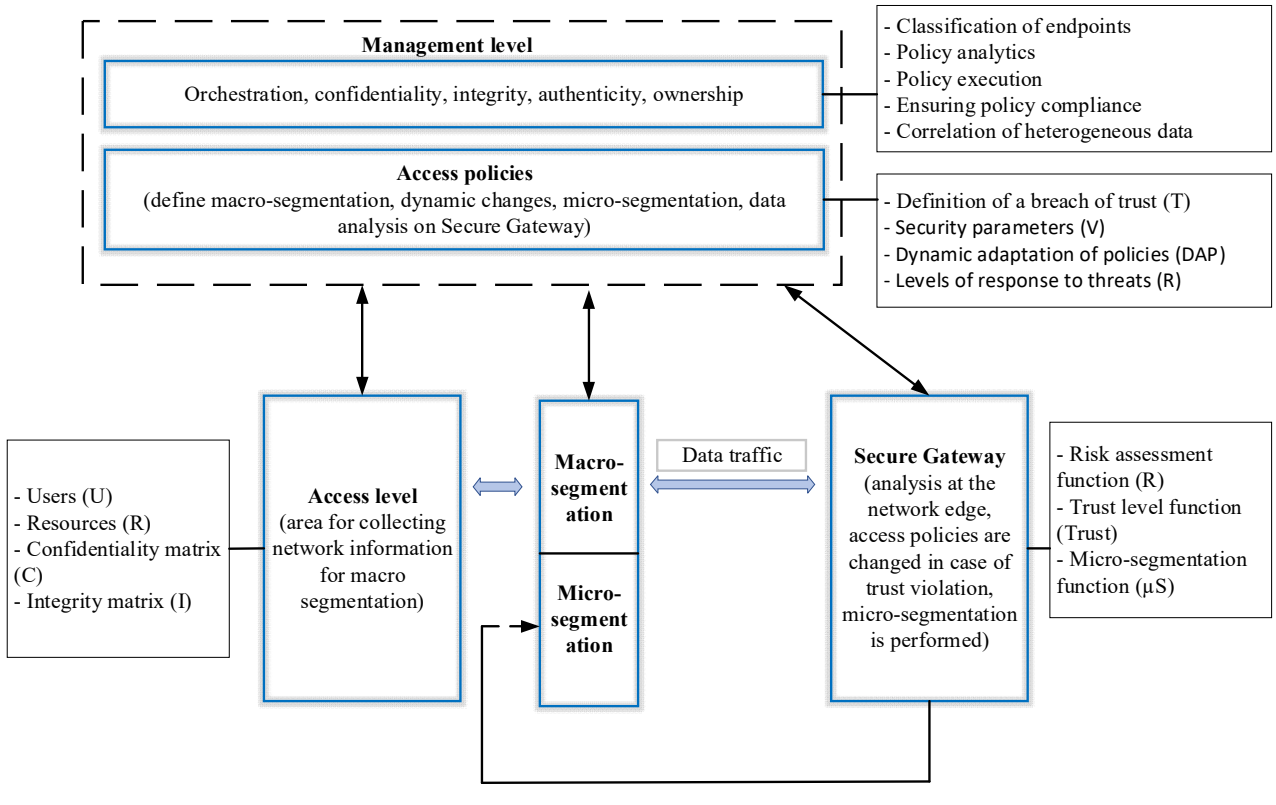

Fig. 2. Structural diagram of traffic segmentation

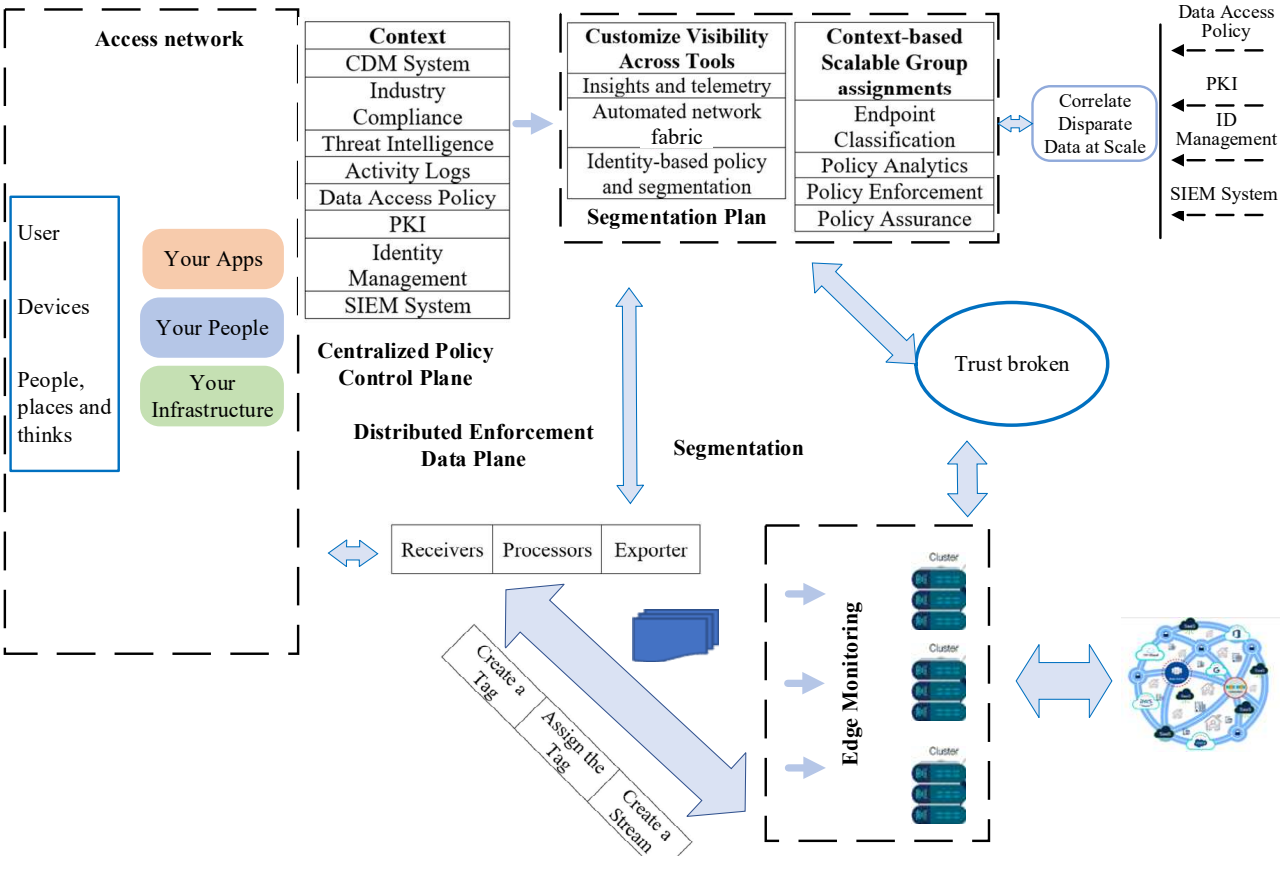

Fig. 3. Traffic security management architecture

Management-level orchestration ensures system-wide security coordination. A "trust broker" plays a key role in trust and security management by analyzing data streams and making decisions about trusting them. Clusters at the edge of the network perform additional traffic monitoring and analysis to ensure that each cluster processes only traffic that conforms to its security policies.

The stages of processing in the current network architecture are presented in the Table 2.

Table 2

Interaction of components

Processing stages	Description
1	2
Access level (Access networks)	Data is collected from users and devices (both mobile and IoT), which may include information about people, locations, and various objects. Data is divided into blocks that are tied to specific applications, people or infrastructure. This is the first step towards traffic identification and classification
Centralized policy control (Centralized Policy Control Plane)	Data flows through a policy control system that uses contextual information such as enterprise mobility management (MDM) systems, security and event management (SIEM) systems, and network automation tools to define and apply security policies. This includes marking data based on its source and intended use, as well as segmentation planning to distribute it across different areas of the network
Distributed policy enforcement (Distributed Enforcement Data Plane)	Data is transmitted over a distributed data plane where security policies are applied according to marking
Dynamic marking and microsegmentation	If trust is broken, the system dynamically changes security policies and marking levels. This allows the system to adjust access and restrictions with more stringent parameters for each subsequent breach of trust, resulting in fine-grained micro-segmentation of traffic. For more detailed dynamic marking and micro-segmentation, monitoring by parameter groups is used. At the same time, due to the complex nature of the traffic, which includes social components of influence, semiotic parameters are introduced for DNS (Domain Name System) attacks [16]
Control at the network edge (Edge monitoring)	Data reaches the edge of the network, passes through an additional level of monitoring and analysis. Clusters process only filtered data streams that match their specifications
Coordination (Trust broker)	The "Trust broker" acts as the coordinator, which coordinates all policy and security management processes. It analyzes data streams, determines trust levels, and makes decisions about marking and applying policies
Automation and integration	Traffic security management processes are supported by an automation and integration system that minimizes manual intervention and increases safety management efficiency. Automation allows to continuously collect and analyze data (insights and telemetry)

Continuation of Table 2

1	2
Warranty and policy application (Policy Assurance, Policy Enforcement)	It is verified that all security measures are applied correctly and that established policies are followed. This includes verifying compliance with policies at the endpoint level (endpoint classification and compliance) and ensuring that the marking and segmentation of traffic complies with established security rules
Dynamic response and adaptation	The system is capable of dynamically responding to threats and changes in the network infrastructure, adapting security policies in real time
Analytics and compliance (Compliance)	Analytical tools such as SIEM (Security Information and Event Management) help correlate and analyze data to identify anomalies and potential threats. Systems also ensure compliance with rules and regulations

This process is cyclical and repeats with refinement of policies and parameters each time a breach of trust is detected, allowing the security system to dynamically adapt to current threats and changes in the environment.

5. Results of a protecting information resources method development for the socio-cyber-physical system based on semiotic analysis

5. 1. Analysis of user and device behavioral data at the access level

To describe the access level, several matrices are introduced that display data for each user and resource:

– the set of users U – the set that contains all users of the system, u_1, u_2, ..., u_m – individual users, m – total number of users;

– the set of resources R – the set containing all the resources to which access is to be controlled, r_1, r_2, ..., r_n – separate resources, n – total number of resources.

For aspects of CIA (Confidentiality, Integrity, Availability), three policy matrices are proposed:

– confidentiality matrix C – a matrix of size $m \times n$, where the element C_{ij} determines the level of privacy that the user u_i needs to access the resource r_j;

– integrity matrix I – a matrix of size $m \times n$, where the element I_{ij} describes the level of integrity that the system must maintain during user u_i interaction with the resource r_j;

– availability matrix A – a matrix of size $m \times n$, where the element A_{ij} indicates the level of resource r_j availability for the user u_i.

To form a complex access policy, combine these three matrices into one matrix of access policies P:

$$P_{ij} = \left(C_{ij}, I_{ij}, A_{ij}\right), \tag{1}$$

where C_{ij} – confidentiality matrix,
I_{ij} – integrity matrix,
A_{ij} – availability matrix.

Thereby, P_{ij} represents a vector or tuple of values that together define a user's u_i access policy to the resource r_j, considering aspects of confidentiality, integrity and availability.

To account for dynamic aspects of security policies, such as access context or changing threat levels, a function f has been added, which will modify or select appropriate values from P depending on additional variables:

$$P_{ij}(t) = f\left(C_{ij}, I_{ij}, A_{ij}, R_i, K_j, T_t\right), \tag{2}$$

where C_{ij} – confidentiality matrix,
I_{ij} – integrity matrix,
A_{ij} – availability matrix.
t – time or session;
R_i – user u_i role;
K_j – resource r_j context;
T_t – threat level at time t.

The function f is defined by the organization's security policies and implemented by a set of rules or algorithms, including decision rules, machine learning, or other techniques.

To extend the access model A, considering the identification of users, devices, applications and behavior history, the following components are introduced:

1. User identification (ID_{user}) is a vector that contains unique identification attributes for a user. Specifically, username, roles, authentication attributes, etc.

2. Device identification (ID_{dev}) is a vector characterizing device j, including its type, operating system, serial number, and other unique identifiers.

3. Application identification (ID_{app}) is a vector that represents application k, including its name, version, access permissions, and any other relevant attributes.

4. Behavioral pistory (H) – activity is logged for each user, device and application H_{user_i}, H_{dev_j}, H_{app_k}, which reflects the history of their interactions with the system and other components.

Incorporating these elements into the access model requires expanding the confidentiality, integrity, and availability matrices:

$$P_{ij} = f\begin{pmatrix} C_{ij}, I_{ij}, A_{ij}, ID_{user}, ID_{dev}, ID_{app_k}, \\ H_{user_i}, H_{dev_j}, H_{app_k}, R_i, K_j \end{pmatrix}, \tag{3}$$

where C_{ij} – confidentiality matrix,
I_{ij} – integrity matrix,
A_{ij} – availability matrix,
R_i – user u_i role,
K_j – resource r_j context,
ID_{user} – identification attributes for the user,
ID_{dev} – device characteristics,
ID_{app} – application characteristics,
H_{user_i} – user activity log,
H_{dev_j} – device activity log,
H_{app_k} – application activity log.

At the same time, the function f becomes more complex because it must integrate different data sources to determine the access level. It uses machine learning techniques and sophisticated rules to adapt access levels based on behavioral patterns and changing context.

To ensure privacy, integrity and availability in a dynamic and adaptive manner, the model uses a risk-based approach, assessing risk based on behavioral history and current identity context:

$$Risk_{ij} = g\left(H_{user_i}, H_{dev_j}, H_{app_k}, ID_{user_i}, ID_{dev_j}, ID_{app_k}\right). \tag{4}$$

Therefore, P_{ij} additionally depends on the risk assessment $Risk_{ij}$, which affects the requirements for the level of authentication and authorization. As a result, the general matrix of access policies P is formed from the complex interaction of all these elements, which allows for dynamic adaptation of security policies to changing conditions and behavior of users, devices and applications.

5. 2. Formation of a two-level dynamic data marking strategy

5. 2. 1. Network macro-segmentation

The network is divided into zones with different levels of trust and security policies. This can be thought of as a function $S: P \rightarrow Z$, where Z – zone set.

For network segmentation, several sets and parameters are defined, which will be used in the model:

1. Device sets D: D_s – a set of devices on a network segment s. These can be client devices, servers, network equipment, etc.

2. Types of applications A: A_s – the set of applications running on the S network segment.

3. Network tree diameter Δ – the maximum number of steps between any two nodes in a network segment ΔS.

4. Properties of information distribution channels C: C_{ij} – properties of the link between devices i and j, such as bandwidth, delay, errors, etc.

5. Security of network segments P: P_s – measure of network segment security S, which may include encryption levels, the presence of firewalls and intrusion detection systems.

Using these parameters, the network segmentation function S can be defined as:

$$S(D, A, \Delta, C, P) = \{s_1, s_2, \ldots, s_n\}, \quad (5)$$

where s_k is a network segment defined in terms of the set of devices, application types, network tree diameter, properties of information distribution channels, and security level;

D – device sets;

A – application types;

Δ – network tree diameter;

C – properties of information dissemination channels;

P – security of network segments.

Moreover, for each segment:

$$s_k = \{D_{s_k}, A_{s_k}, \Delta_{s_k}, C_{s_k}, P_{s_k}\}. \quad (6)$$

The evaluation function E, which estimates the appropriate security level for each segment based on its parameters, has the following form:

$$E(D_{s_k}, A_{s_k}, \Delta_{s_k}, C_{s_k}) = P_{s_k}. \quad (7)$$

Network segmentation is dynamic. It adapts to changes in network traffic, application behavior, and threat levels.

5. 2. 2. Data block marking

Each data block is marked according to zone policies:

$$M: D \times Z \rightarrow T, \quad (8)$$

where D – data blocks; T – marking tags.

Data block marking M can be represented as a function that integrates aspects to generate a mark:

$$M(d) = f\begin{pmatrix} CDM_d, IC_d, TI_d, AL_d, DAP_d, \\ PKI_d, IDM_d, SIEM_d \end{pmatrix}, \quad (9)$$

where CDM_d – continuous security monitoring component;

d – data block;

IC_d – the degree of compliance of the data block d with industry standards;

TI_d – a metric that assesses the level of threat associated with the data d;

AL_d – historical data about interactions with the data block d;

DAP_d – describes the policies that determine how and by whom the d data may be used;

PKI_d – denotes the use of a public key infrastructure to ensure d data security;

IDM_d – component of identification and authentication mechanisms related to d data;

$SIEM_d$ – impact of the SIEM system on the d data.

The function f combines information from different sources and generates a unique mark M_d for each block of data. This M_d tag includes information that will be used to make decisions in further stages of data processing and protection, such as network segmentation, application of security policies, and analysis at the network edge.

The following components have been introduced to mark M data blocks:

1. CDM (Continuous Diagnostics and Mitigation) system. This system provides continuous security monitoring, which can be defined as CDM_d, where d – is a data block.

2. Compliance with industry standards (Industry Compliance). IC_d – a boolean value or set of values that reflect the extent to which the data block d conforms to industry standards.

3. Intelligent threat analysis (Threat Intelligence). TI_d – a metric that assesses the level of threat associated with the data d.

4. Activity logs. AL_d – historical data about interactions with the data block d.

5. Data Access Policy. DAP_d describes the policies that determine how and by whom the d data may be used.

6. Public key infrastructure (PKI). PKI_d denotes the use of a public key infrastructure to secure data d, including encryption and digital signatures.

7. Identification management (ID Management). IDM_d refers to the identification and authentication mechanisms associated with the data d.

8. Information security management system (SIEM System). $SIEM_d$ reflects the impact of a SIEM system on data d, such as security monitoring and analysis.

5. 2. 3. Analysis at the edge of the Secure Gateway

At the edge of the network, blocks of information are again analyzed and sent to the appropriate clusters, $G: T'' \rightarrow C$, where C – a set of clusters at the edge of the network.

At the same time, it is possible to define the functions of risk assessment and the level of trust:

1. The risk assessment function R estimates the risk level r for the traffic flow x:

$$R(x, NL, RL, DNSL, WL, VoL, BP, IoTA, Perf, H) = r, \quad (10)$$

where x – traffic flow;

NL – network level parameters;

RL – routing level parameters;
DNSL – DNS level parameters;
WL – web level parameters;
VoL – voice traffic level parameters;
BP – biographical properties of domain traffic;
IoTA – metrics related to attacks on IoT devices;
Perf – network performance under normal conditions and during DDoS attacks;
H – historical data.

2. The Trust level function determines the level of trust t based on risk assessment and historical data:

$$Trust(r,H)=t, \quad (11)$$

where r – risk level; H – historical data.

3. Microsegmentation function μS assigns new security tags T'' based on previous tags T', risk r, and level of trust t:

$$\mu S(T',r,t)=T'', \quad (12)$$

where T' – initial security tags before breach of trust;
r – risk level;
t – level of trust.

When trust is violated, the system updates security policies:

$$Pg^{'}=Update(Pg,Corr(DAP,PKI,IDM,SIEM)), \quad (13)$$

where *Update* – policy update function;
Corr – data correlation function;
Pg – current security policy;
DAP – data access policy;
PKI – public key infrastructure;
IDM – identity management data;
SIEM – information security management system data;
Pg' – updated security policy.

Traffic segmentation is updated as follows:

$$S'(x,Pg')=Segment(x,Pg'), \quad (14)$$

where *Segment* – a segmentation function that assigns traffic flows to new network segments;
x – traffic flow;
Pg' – updated security policy that applies to traffic.

To eliminate implicit trust in the network, the system must require continuous verification and updating of trust levels for all devices and traffic in real time:

$$T''(x)= =\mu S\begin{pmatrix} T'(x),R\begin{pmatrix} x,NL(x),RL(x),DNSL(x),\\ WL(x),VoL(x),BP(x),\\ IoTA(x),Perf(x),H(x) \end{pmatrix},\\ Trust(R(x,\ldots),H(x)) \end{pmatrix}, \quad (15)$$

where T'' – final security tags after micro-segmentation that reflect the adaptive security measures taken in response to the current level of risk and trust for each traffic flow;
x – a traffic flow or block of data to be made and segmented;
$NL(x)$, $RL(x)$, $DNSL(x)$, $WL(x)$, $VoL(x)$ – functions that extract the appropriate parameters for flow x from each level of the network architecture;
$BP(x)$, $IoTA(x)$, $Perf(x)$ – functions that analyze flow x for domain traffic biographical properties, IoT attack signatures, and performance metrics, respectively;
$H(x)$ – historical data and behavioral patterns associated with flow x;
$T'(x)$ – the initial security tags assigned to flow x after the initial network segmentation.

5. 2. 4. Microsegmentation

For each violation of trust, a narrowed set of parameters V' is used and perform microsegmentation. This can be modeled as a function μS: $T'\times V'\rightarrow T''$, where T'' – tags after microsegmentation.

Micro-segmentation and additional marking of traffic, which is carried out after the initial segmentation of the network, is described using matrices or functions that take into account different levels of the network infrastructure and parameters of semiotic analysis.

Microsegmentation and marking can be represented by a function μS, which accepts tags after primary segmentation T', parameters of different levels *NL, RL, DNSL, WL, VoL* and parameters of semiotic analysis *BP, IoTA, Perf*:

$$\mu S(T',NL,RL,DNSL,WL,VoL,BP,IoTA,Perf)=T'', \quad (16)$$

where T'' – a new set of tags after micro-segmentation. It details the characteristics and security requirements for each block of data or traffic, encryption for certain types of traffic or applying stricter filtering rules for traffic from untrusted sources are key elements of micro-segmentation;
NL – a set of parameters describing network-level characteristics;
RL – a set of parameters describing routing within a network;
DNSL – parameters related to DNS;
WL – parameters related to web traffic;
VoL – parameters related to voice traffic;
BP – biographical properties of domain traffic;
IoTA – parameters that identify potential attacks on IoT devices;
Perf – performance measurement.

The micro-segmentation feature includes a series of heuristics or algorithmic rules to determine the best way to isolate and protect traffic based on its characteristics.

If IoTA or Perf indicates an attack, the traffic is directed to a high security zone (`high_sec_zone`) that requires special protection due to suspected attack or high risk.

If the BP confirms the trust of the domain, the traffic is directed to the standard security zone (`standard_zone`) designated for trusted traffic.

If the WL or VoL does not comply with security policies, the traffic is directed to a restricted zone (`restricted_zone`) where access is restricted due to non-compliance with certain security policies:

1. Network level (NL). *NL* – a set of parameters describing network-level characteristics such as IP addresses, subnet masks, VLAN IDs, and so on.

2. Routing level (RL). *RL* – a set of parameters describing routing within a network, including routing tables, routing policies, protocols, etc.

3. DNS-level (DNSL). *DNSL* – DNS-related settings such as DNS records, DNSSEC settings, etc.

4. Web-level (WL). *WL* – parameters related to web traffic, including HTTP headers, cookies, SSL/TLS certificates, etc.

5. Voice level (VoL). *VoL* – parameters related to voice traffic, such as SIP headers, RTP streams, etc.

6. Semiotic analysis:

– domain traffic biographical properties (BP): includes domain history and reputation analysis;

– IoT attacks (IoTA): refers to parameters that identify potential attacks on IoT devices;

– performance measurement (Perf): Perf includes metrics such as latency, data rate, etc. measured during normal authentications or DDoS attacks.

5. 3. Development of a method of protection of mixed content of socio-cyberphysical system information based on semiotic analysis

5. 3. 1. Security policies

Security policies are applied during macro-segmentation, micro-segmentation, in determining the level of trust violation. They can change dynamically at the management level. Fine-grained security policies are applied during micro-segmentation, assigned to applications from the data center, down to the workload level, as well as devices. This means that security policies can be synchronized with a virtual network, virtual machine, operating system, or other virtual security targets.

Security policies not only adapt to the current breach, but also consider the potential risk to determine appropriate response levels and strengthen security measures.

If trust is broken, dynamic security policies are applied. A policy is a set of rules $B: T \times V \to T'$, where V – security parameters, and T' – new post-breach tags that adapt in response to trust breaches. These rules are formalized as functions that define new security measures based on the current security context and breach history:

1. Defining breach of trust *T*. $T(d, h)$ – a function that evaluates whether trust has been violated for a block of data d given the history h (which includes activity logs, threat analysis, and previous violations).

2. A set of security parameters *V*. V – a parameter vector that includes encryption level, access control types, authentication requirements.

3. Dynamic adaptation of policies DAP. Dynamic adaptation of policies $DAP(d, v, t)$ is a function that modifies the security parameters v in response to a trust violation t for data d.

4. Threat response levels *R*. Threat response levels $R(v, t)$ can be defined as a set of strategies or procedures that are activated depending on the threat level t and the current security parameters v.

With these components in mind, dynamic security policies are presented as follows.

Threat response levels R are integrated into the model:

$$B'(d,h,v) = R\left(DAP\left(d,v,T(d,h)\right),T(d,h)\right), \quad (17)$$

where d – data or traffic;

h – historical data related to d;

v – the current set of security parameters;

T – function that determines whether the trust has been violated returns a Boolean value (true or false);

DAP – a dynamic policy adaptation function that updates v parameters in response to trust violations.

The DAP function is responsible for changing the security parameters in case the trust in the data d has been compromised. Otherwise, the parameters remain unchanged. This forms the basis for the dynamic adaptation of security policies in a zero trust model, where security is constantly re-evaluated and adapted to the current context.

5. 3. 2. Coordination at the management level

Coordination at the management level (Control Plane) takes into account the classification of endpoints, policy analytics, policy implementation and policy compliance, as well as the correlation of disparate data (Fig. 3):

1. Endpoint Classification. Each endpoint e in the network is classified based on a set of attributes A_t, which include device type, operating system, software version, and other characteristics.

2. Policy Analytics. For each set of policies, P_g, the analytical process evaluates the compliance of each endpoint with existing policies and generates a set of rules, $R_{e.g}$, that determine which policies are applicable to each endpoint.

3. Policy Enforcement. Based on the generated rules $R_{e.g}$, the system applies policies to the endpoints described by the function $E(e, R_{e.g})$. This results in a set of Act_e actions that must be performed to ensure policy compliance.

4. Policy Assurance. To check and confirm that policies have been applied correctly, the *Assure* (e, Act_e) function is used, which returns a binary result $V_e^{'}$ (checked/not checked).

5. Correlate Disparate Data. When a trust breach occurs, the system correlates data from various sources, such as data access policies, PKI, identity management, and SIEM systems, to determine what changes are needed in security policies. This is represented by a function:

$$Corr\left(DAP, PKI, IDM, SIEM, H\right), \quad (18)$$

where H – historical data related to endpoints and traffic. This function evaluates inconsistencies and anomalies in data to identify potential threats or breaches,

DAP – policies that determine how and by whom the data may be used,

PKI – using public key infrastructure to ensure data security,

IDM – a component of data-related identification and authentication mechanisms,

SIEM – impact of the SIEM system on the data:

$$Corr\left(DAP, PKI, IDM, SIEM, H\right) = \Delta, \quad (19)$$

where Δ – a set of changes or fixes that need to be applied to the current network security policy or configuration.

6. Dynamic Policy Adaptation. When trust is broken, the system must adapt security policies based on the output of the correlation function Δ. This leads to updates or changes to $R_{e.g}$ rules, which then affect policy enforcement and compliance:

$$\begin{aligned} R_{e.g}^{'} &= Adapt\left(R_{e.g}, \Delta\right), \\ Act_e^{'} &= E\left(e, R_{e.g}^{'}\right), \\ V_e^{'} &= Assure\left(e, Act_e^{'}\right), \end{aligned} \quad (20)$$

where $R_{e.g}^{'}$ – updated rules after adapting policies;

$Act_e^{'}$ – updated actions to be taken in response to adapted policies;

$V_e^{'}$ – the results of checking updated actions;

Δ – a set of changes or fixes that need to be applied to the current network security policy or configuration;

$R_{e.g}$ – a set of generated rules;

e – an endpoint in a network that is classified based on a set of attributes A_e.

5. 4. Verification of the developed mixed content protection method of socio-cyber-physical system information based on semiotic analysis

Types of network traffic data are quite diverse and it is very difficult to take into account and combine all aspects. Therefore, the DNS protocol was chosen to simulate the developed method, because in addition to classic "man-in-the-middle" attacks, privacy violation attacks, it takes into account influence attacks with information manipulation.

DNS (Domain Name System) is one of the first and most vulnerable network protocols, with several security weaknesses that have been exploited repeatedly by attackers over the years. DNS abuse has always been a major concern for cybersecurity researchers. However, ensuring the security and privacy of DNS queries and responses remains a challenge, as attackers use sophisticated attack techniques to steal data on the fly.

Modeling was done using Splunk, Python, CVE security vulnerability database (CVE), Automated Indicator Sharing (AIS) and Common Vulnerability Scoring System (CVSS). Several modules and add-ons from Splunk Enterprise Security (ES) were used to model and develop solutions in the field of cyber security using Splunk. The Malicious DNS and Attacks data set (BCCC-CIC-Bell-DNS-2024), obtained as a result of an experiment conducted within the framework of the Canadian Cyber Security Institute (CIC) Project, funded by the Canadian Internet Registration Authority (CIRA) was used as the source data [19]. This suite uses the DNS over HTTPS (DoH) protocol in RFC8484, which was introduced to address some of the privacy and data manipulation vulnerabilities of DNS, but is still vulnerable in enterprise network traffic. The set presents test and evaluation data for DoH traffic in covert channels and tunnels. The data covers both useful and malicious DoH traffic in a two-level approach to detecting and characterizing DoH traffic using a time series classifier [20, 21].

Each group of current data parameters (lexical, DNS statistical, parameters from third parties) affects the model differently:

– lexical parameters (F1–F14) affect initial domain classification and initial segmentation;

– DNS statistics options (F15–F21) provide dynamic segmentation updates based ON current network activity and statistics;

– Third-party parameters (F22–F32) are used for deeper analysis and more stringent micro-segmentation if the suspicion of malicious activity increases.

To model network security parameters and their presentation, the evaluation criteria described above were used, which were grouped by categories corresponding to the level of segmentation detail (Table 3) [19]. Table 3 contains the main parameters, their description and corresponding groups.

Table 3

Basic parameters for modeling network security

Parameter	Description	Segmentation category	Parameter group
1	2	3	4
F1: Subdomain	Whether the domain has a subdomain or not	Lexical segmentation	Lexical parameters
F2: TLD	The top level of the domain	Lexical segmentation	Lexical parameters
F3: SLD	The second level of the domain	Lexical segmentation	Lexical parameters
F4: Len	Length of domain and subdomain	Lexical segmentation	Lexical parameters
F5: Numeric percentage	The percentage of numbers in the domain and subdomain	Lexical segmentation	Lexical parameters
F6: Character distribution	Distribution of symbols in the domain	Lexical segmentation	Lexical parameters
F7: Entropy	Entropy of letter distribution	Lexical segmentation	Lexical parameters
F8: 1-gram	1-gram domain at the letter level	Lexical segmentation	Lexical parameters
F9: 2-gram	2-gram domain at the letter level	Lexical segmentation	Lexical parameters
F10: 3-gram	3-gram domain at the letter level	Lexical segmentation	Lexical parameters
F11: Longest word	The longest meaningful word in the second level of the domain (SLD)	Lexical segmentation	Lexical parameters
F12: Distance from bad words	Average distance to "bad" words	Lexical segmentation	Lexical parameters
F13: Typos	Typosquatting (mistakes in the domain set)	Lexical segmentation	Lexical parameters
F14: Obfuscation	Maximum value for URL obfuscation	Dynamic segmentation	Lexical parameters
F15: Unique country	The number of unique countries in the time window τ	Dynamic segmentation	DNS statistical parameters
F16: Unique ASN	The number of unique ASN values in the time window τ	Dynamic segmentation	DNS statistical parameters
F17: Unique TTL	The number of unique TTL values in the time window τ	Dynamic segmentation	DNS statistical parameters
F18: Unique IP	The number of unique IP addresses in the time window τ	Dynamic segmentation	DNS statistical parameters
F19: Unique domain	The number of unique domains in the time window τ	Dynamic segmentation	DNS statistical parameters
F20: TTL mean	Average TTL value in time window τ	Dynamic segmentation	DNS statistical parameters
F21: TTL variance	Dispersion of TTL values in the time window τ	Dynamic segmentation	DNS statistical parameters
F22: Domain name	Domain name	Microsegmentation	Parameters from third parties
F23: Registrar	The domain registrar	Microsegmentation	Parameters from third parties
F24: Registrant name	The name the domain is registered to	Microsegmentation	Parameters from third parties

Continuation of Table 3

1	2	3	4
F25: Creation date time	Date and time of domain creation	Microsegmentation	Parameters from third parties
F26: Emails	Emails associated with the domain	Microsegmentation	Parameters from third parties
F27: Domain age	Domain age	Microsegmentation	Parameters from third parties
F28: Organization	The organization with which the domain is associated	Microsegmentation	Parameters from third parties
F29: State	The state in which the main office is located	Microsegmentation	Parameters from third parties
F30: Country	The country in which the main branch is located	Microsegmentation	Parameters from third parties
F31: Name server count	The total number of name servers associated with the domain	Microsegmentation	Parameters from third parties
F32: Alexa rank	Domain ranking according to Alexa	Microsegmentation	Parameters from third parties

Segmentation categories:

– lexical segmentation: used to initially classify and segment domains based on their textual characteristics;

– dynamic segmentation: used to adaptively update segmentation based on real-time analysis of DNS statistical characteristics;

– micro-segmentation: used for detailed analysis and segmentation of domains based on parameters related to domain registration and history.

The purpose of this simulation is to determine data in data flows according to their characteristics and risks at the level of Edge Monitoring and Trust broken blocks (Fig. 3). This is achieved through risk assessment, trust and micro-segmentation of data flows within the network. Using a variety of parameters to assess each flow provides a deep understanding of potential threats and vulnerabilities. Options include:

– *NL* (*NetworkLevel*): IP addresses, subnet masks, VLAN ID;

– *RL* (*Routing Level*): Routing tables, routing policies;

– *DNSL* (*DNS Level*): DNS records, DNSSEC settings;

– *WL* (*Web Level*): HTTP headers, cookies, SSL/TLS certificates;

– *VoL* (*Voice Level*): SIP headers, RTP streams;

– *BP* (*Biographical Properties*): Domain history and reputation;

– *IoTA* (*IoT Attacks*): Parameters identifying potential IoT attacks;

– *Perf* (*Performance Measurements*): Latency, data transfer rate.

Before implementing evaluation algorithms as part of a Zero Trust architecture, it is important to understand how each component of the infrastructure interacts and affects the overall security of the system. Key elements such as users, devices, data and networks are constantly monitored and analyzed to identify potential risks and vulnerabilities (Table 4). Using a comprehensive approach to authentication, data encryption, application protection and threat response automation allows to create a more reliable protection system. With this in mind, the following evaluation algorithms are implemented for effective security management:

1. Risk Rating Function (*R*): a combination of the above parameters to determine the risk level of each flow.

2. Trust Level Function (*Trust*): determination of trust based on risk assessment and historical data.

3. Micro-segmentation (**μS**) feature: assign new security tags based on existing tags, risk, and trust level.

Simulation includes steps that simulate real operating conditions in the network:

– flow definition: classification of incoming data streams based on their parameters;

– risk assessment and trust: using machine learning algorithms for flow analysis and risk assessment;

– micro-segmentation: applying security policies to isolate and manage flows based on their risk and trust assessments;

– security policy adaptation: real-time security policy updates based on analysis and micro-segmentation.

Table 4

Parameters for modeling network security

Parameter	Description	Segmentation category	Parameter group
SSL Certificate Validity	SSL certificate validity	Microsegmentation	Certification parameters
SSL Certificate Authority	Authoritativeness of the certification center	Microsegmentation	Certification parameters
HTTP Headers Security	Security of HTTP headers	Dynamic segmentation	Web-parameters
Content Management System	Content Management System (CMS) used	Microsegmentation	Web-parameters
Server Software	Server software	Microsegmentation	Infrastructure parameters
IP Blacklist Status	Blacklisted IP address status	Dynamic segmentation	Network parameters
Geo-IP Location	Geographical location of the IP address	Microsegmentation	Network parameters
Traffic Volume	Traffic volume	Dynamic segmentation	Network performance
Traffic Patterns	Traffic patterns	Dynamic segmentation	Network performance
Behavior Anomalies	Behavioral abnormalities	Dynamic segmentation	Behavior analysis
API Call Patterns	API call patterns	Microsegmentation	Program parameters
Cloud Service Provider	Cloud service provider	Microsegmentation	Infrastructure parameters
Data Breach History	History of data leaks	Microsegmentation	Historical parameters
Patching Cadence	Frequency of applying patches	Microsegmentation	Operational parameters
Encryption Strength	Encryption strength	Microsegmentation	Cryptographic parameters
Authentication Methods	Authentication methods	Microsegmentation	Access security
Compliance Status	Regulatory compliance status	Microsegmentation	Legal parameters

For the purpose of modeling, several parameters that are important for the analysis of network traffic and security aspects were selected to optimize the calculations (Table 5):

1. FlowSentRate – the rate at which bytes are sent. Load and potentially suspicious activity if abnormally high.

2. FlowReceivedRate – byte acquisition rate. Detection of anomalies in the speed of data acquisition.

3. PacketLengthMean – average length of packets. Traffic type and potentially malicious payload.

4. PacketLengthVariance – packet size variation.

5. ResponseTimeTimeMean – average response time. Latencies that indicate network issues or heavy processing tasks such as encryption/decryption.

6. ResponseTimeTimeVariance – response time variation.

Table 5

Network traffic and security analysis parameters

No.	Flow Sent Rate	Flow Received Rate	Packet Length Mean	Packet Length Variance	Response Time Time Mean	Response Time Time Variance
0	9,972.59	6,123.52	92.00	338.00	0.0171	0.0000
1	2,675.94	5,943.62	257.19	310,613.30	0.0107	0.000084
2	5,141.87	6,587.54	132.26	13,602.69	0.0109	0.000064
3	1,545.86	3,976.32	347.50	600,872.00	0.0798	0.001276
4	2,009.23	3,959.50	258.67	208,270.50	0.0547	0.001192
5	3,227.89	26,297.72	761.03	1,355,542.00	0.0155	0.000180
6	8,619.23	145,062.65	1,022.77	3,033,225.00	0.0082	0.000097
7	5,186.67	13,489.18	461.84	827,630.00	0.0682	0.000932
8	881.53	1,509.46	120.79	8,124.75	0.0132	0.000045
9	498.23	5,015.77	652.18	983,613.50	0.0127	0.000185
10	174.54	5,058.28	1,405.13	3,457,346.00	0.0066	0.000165
11	1,083.43	20,728.74	829.81	1,317,212.00	0.0073	0.000120
12	475.59	1,634.45	344.07	564,384.10	0.0163	0.000191
13	541.39	16,599.55	1,220.16	2,282,238.00	0.0048	0.000097
14	5,567.72	17,785.58	339.53	710,961.40	0.0293	0.000171
15	6,844.65	4,202.86	92.00	338.00	0.0250	0.0000
16	11,053.65	6,787.33	92.00	338.00	0.0155	0.0000
17	2,888.31	7,637.17	327.12	510,841.80	0.0877	0.015396
18	7,153.86	149,600.44	1,176.24	2,168,801.00	0.0036	0.000264
19	14,743.17	230,603.47	961.62	1,420,894.00	0.0057	0.000113

To assess the level of trust for network traffic, a risk score and historical data were selected for each traffic flow.

To calculate the confidence level:

1. Risk Rating (*R*) function: this function estimates the risk level for each traffic flow based on various network parameters (10).

2. Trust Level Function (*Trust*): this function calculates the level of trust based on the level of risk obtained from the above assessment and the historical data associated with each traffic flow.

A simplified linear relationship between these parameters and risk, as well as between risk and trust, was assumed.

Assumption:

– risk score (*R*): calculated as a weighted sum of the normalized values of the selected parameters;

– trust level (*Trust*): inversely proportional to the risk indicator, adjusted for historical reliability indicators.

Used parameters:

– FlowSentRate and FlowReceivedRate to display network performance;

– PacketLengthVariance to indicate variability and potential anomalies in traffic;

– ResponseTimeTimeVariance to capture response time inconsistencies that may indicate potential threats or system overload.

For this example, it was assumed that each of these parameters equally affect the risk assessment, and the reliability of historical data has a simple positive effect on the trust assessment (Fig. 4).

The graphs present numerical values of risk scores and corresponding confidence levels for a set of network endpoints where:

– risk score: values range from approximately 0 to 1. A higher score indicates a higher perceived risk associated with that particular endpoint or activity based on network parameters;

– confidence level: values range from approximately 0 to 1, where a higher value indicates a higher level of confidence. Confidence levels are inversely proportional to risk estimates, potentially governed by the historical reliability of the endpoints. Confidence level calculations include historical data for buffering.

The figure quantifies the risk associated with each endpoint, providing a clear metric for effectively evaluating security postures. Numerical trust metrics are shown, allowing to understand which areas of the network or which activities are considered more reliable based on past and current data. Correlation between risk assessments and confidence levels is considered in the dynamic adjustment of security policies.

Security policies are applied during macro-segmentation, micro-segmentation, in determining the level of trust violation. They change dynamically at the management level. Detailed security policies are applied during micro-segmentation. When trust is broken, the system must adapt security policies based on the output of the correlation function Δ.

A set of changes or adjustments (Δ):

– updating endpoint attributes;

– integration of advanced analytics;

– adaptive law enforcement mechanisms;

– improved verification methods;

– advanced data correlation;

– quick policy adaptation.

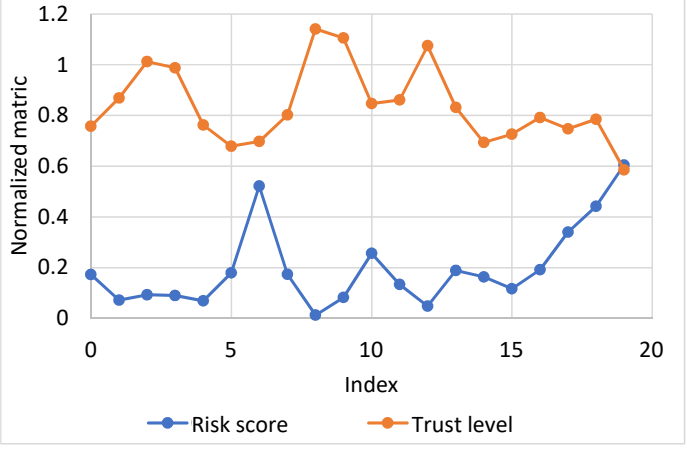

Fig. 4. Calculation of risk assessment and confidence level

Table 6 presents the simulation results that reflect the impact of the proposed changes and security policy configurations on 20 endpoints (Table 6).

Values range from 0 to 100, where higher values indicate better performance or compliance.

Table 6 represents the improvement in security measures with the current model, where each row is an endpoint showing potential improvements in several dimensions due to the changes made.

Endpoint Compliance Score (%) reflects the percentage to which endpoints comply with established security policies.

Policy Update Frequency – frequency of updates per year. Shows how many times per year security policies are reviewed and updated. Higher values indicate more frequent updates, which can help improve adaptability to changing threats.

Table 6

Simulation results

Index	Endpoint Compliance Score	Policy Update Frequency	Anomaly Detection Rate	Enforcement Success Rate	Assurance Verification Rate	Data Integration Level
0	81.25	41.98	61.49	64.65	77.21	74.36
1	93.00	50.22	58.22	77.34	71.87	76.85
2	84.47	37.88	49.66	70.85	70.41	72.69
3	89.74	45.96	61.43	84.92	93.91	56.81
4	79.33	32.45	56.98	63.69	77.78	53.72
5	76.91	46.25	58.43	71.44	81.17	74.40
6	94.18	49.89	53.72	73.68	80.12	70.74
7	83.82	35.71	51.12	67.19	75.26	56.84
8	89.83	55.29	48.81	76.58	86.45	57.05
9	77.67	40.57	60.21	75.04	87.74	62.86
10	85.60	35.81	50.22	79.65	81.91	62.48
11	78.80	36.43	55.79	72.54	77.87	61.94
12	75.50	36.41	55.03	68.47	98.37	69.76
13	81.13	47.02	58.92	66.96	72.38	65.18
14	84.42	33.75	49.51	67.12	78.41	69.94
15	76.63	56.87	66.27	80.48	81.13	53.79
16	73.85	48.72	57.91	81.98	94.25	72.66
17	81.16	32.99	44.79	70.67	93.77	57.44
18	83.31	39.06	63.09	71.77	72.58	76.21
19	88.41	41.15	57.80	70.11	84.60	63.08

Anomaly Detection Rate (%) The percentage of anomalies that were detected. Higher values indicate a more effective system for detecting potential threats.

Enforcement Success Rate (%). Shows the effectiveness of the implementation and application of security policies.

Assurance Verification Rate (%). The percentage of cases in which the correct application of policies was confirmed.

Data Integration Level (%). Indicates the depth and effectiveness of integrating data from various sources for threat analysis.

Policy and configuration changes can improve the overall security posture of a network, reduce risks, and improve security compliance, making systems more reliable and capable of resisting modern threats (Table 7).

Table 7 provides a quantitative analysis of the impact of improvements with the implementation of the following model and shows that the developed method of a comprehensive security system not only increases the level of information resources protection, but also provides flexibility in managing cyber security.

Table 7

Quantitative improvement indicators

Indicator	Before implementation	After implementation	Improvement
Level of policy compliance (%)	70	95	+36
Incident response time (hours)	24	3	–88
Detection of complex threats (%)	40	55	+15
User trust level (%)	60	85	+42

The simulation did not take into account unpredictable external attacks or global changes in the network environment, local specific aspects of the network infrastructure, complex aspects such as cryptographic indicators or evaluation of the real performance of the SIEM system.

6. Discussion of the information resources protection method development results of the socio-cyber-physical system based on semiotic analysis

During the research, it was found that the modern information architecture of corporate networks requires a comprehensive approach to information management and its segmentation. The results of this study indicate that the method of protecting information resources based on CISA's Zero Trust Maturity Model can effectively analyze in detail the interactions between applications, users and network infrastructure. This is critical to ensure a high level of security. A method of targeted traffic segmentation was proposed. The introduction of an improved strategy of macro-segmentation and micro-segmentation of network traffic in the corporate network, which includes the use of semiotic analysis parameters – syntactic, semantic and pragmatic, significantly improved the security measures of information resources (Table 7).

Analysis based on user and device behavioral data allows network traffic to be segmented into macro segments, such as separate areas for IoT or enterprise applications. The use of semiotic parameters allows for a deeper analysis of the meaning of data, its context, social aspects and relationships between traffic components. This provides a comprehensive understanding of how information is used and how it can affect network security. The implementation of two-level dynamic data marking provides the ability to adapt protection mechanisms in real time, responding to changes in network activity and new threats.

Previous studies have focused exclusively on the general principles of building multi-contour security systems and levels of threat classification in socio-cyber-physical systems [9, 22]. Unlike them, this study combines the structure of the building with the principle of processing network traffic in the corporate network.

For example, in contrast to the work [18], which developed the approach of applying semiotic parameters of analysis – syntactic, semantic and pragmatic, this research

extends its application to the real structure of the network. The innovative use of macro-segmentation and micro-segmentation of network traffic based on CISA's Zero Trust Maturity model made it possible to improve the performance of information resource security measures.

One limitation of this study is its reliance on the Malicious DNS and Attacks BCCC-CIC-Bell-DNS-2024 dataset, which may not represent all types of social media content. The proposed system may not reflect the mass influx of content from real social networks.

The main drawback of the developed information protection system in the corporate network based on the strategy of zero trust is the dependence on the quality and completeness of analytical data used for traffic verification and segmentation. Although the strategy involves analyzing the behavior of users, devices and traffic, it may face limitations in the variety of data, in the processing of large volumes of data, with the risks of marking errors.

This research can be extended by expanding the data set to include more diverse social media content. Further research could also explore the integration of additional contextual information, such as user browsing history and real-time behavior, to improve the accuracy and relevance of content analysis.

7. Conclusions

1. Behavioral data, context of resource use, and social aspects of interaction between users and devices at the access level were analyzed. The modern information architecture of corporate networks requires a comprehensive approach to information management and its segmentation. Using CISA's Zero Trust Maturity Model, a targeted traffic segmentation method was proposed. This method allows detailed analysis of interactions between applications, users and network infrastructure, which increases the level of detection of complex threats by 15 %.

2. Two-level dynamic data marking is proposed to improve analysis while protecting information resources. The application of marking and micro-segmentation means that any traffic or device on the network must constantly prove its security. In particular, a semiotic analysis of information with its subsequent marking is proposed, which includes biographical characteristics of domains, detection of IoT attacks, and assessment of network behavior during endpoint authentication and DDoS attacks. This aspect makes the system more flexible and adaptable to modern cyber security challenges. Coordination of the system with the help of dynamic access policies allows it to continuously analyze and respond to changes in user behavior and potential threats, thereby providing a high level of protection. Micro-segmentation of the network is carried out taking into account the technology of transmission in the network based on the results of semiotic analysis. At the same time, it ensures the efficiency of information and traffic separation in accordance with security needs. Thus, the transmitted information passes through several levels of verification and segmentation, where each stage adds an additional level of protection.

An improved strategy of two-level dynamic data marking of network traffic in the corporate network was introduced, in which the semiotic parameters of the analysis – syntactic, semantic and pragmatic – are applied. This strategy aims to improve security level analysis, which reduces incident response time by 88 %.

3. A method of protecting the mixed content of information of the socio-cyber-physical system based on semiotic analysis has been developed, which not only increases the level of protection of information resources, but also provides flexibility in managing cyber security. This allows not only to detect and respond to threats in real time, but also to adapt security policies according to the dynamics of user behavior and general security conditions. The use of advanced analytical tools and fuzzy logic allows to detect and respond to complex attacks that traditional security systems may not notice. Implementation of a clear and transparent security system based on verification and control strengthens users' trust in the information security of the corporate network by 42 %.

4. The verification of the developed method of protection of the mixed content of information of the socio-cyber-physical system was carried out on the basis of semiotic analysis, which was based on the strategy of zero trust and the use of modern analytical technologies. In the course of the simulation, the qualitative characteristics of the system were determined, which allow detecting and responding to cyber threats much faster, as well as quantitative indicators that contribute to a prompter response to security incidents. Thanks to the use of analytical modeling and fuzzy logic, improvements in network traffic segmentation and identification of anomalous behavioral patterns have been achieved. This ensured an increase in the overall effectiveness of the information protection system and strengthened user confidence in cyber security measures in the network.

Thus, the integration of analytical methods and modern technologies in the security strategy not only provides protection against current threats, but also creates the basis for an adaptive and sustainable cyber defense system capable of facing future challenges in the field of cyber security.

Acknowledgements

The authors express great gratitude (posthumously) to Professor, Doctor of Technical Sciences Oleksandr Serkov, who made a great contribution to the research.

Conflict of interest

The authors declare that they have no conflict of interest in relation to this study, including financial, personal, authorship, or any other, that could affect the study and its results presented in this article.

Financing

The research was carried out without financial support.

Data availability

The manuscript has no associated data.

Using artificial intelligence tools

The authors confirm that they did not use artificial intelligence technologies when creating the presented work.

References

1. NIST Special Publication 800-207. Zero Trust Architecture (2020). U.S. Department of Commerce. National Institute of Standards and Technology Special Publication 800-207 Natl. Inst. Stand. Technol. Spec. Publ. 800-207, 59. Available at: https://doi.org/10.6028/NIST.SP.800-207
2. Jammine, A., Serkov, A., Lazurenko, B., Nait-Abdesselam, F. (2023). The Order of Formation of Information Signals in IIoT. IJCSNS International Journal of Computer Science and Network Security, 23 (3), 139–143. https://doi.org/10.22937/IJCSNS.2023.23.3.14
3. Standard ISO/IEC 27032:2023 (2023). Cybersecurity. Guidelines for Internet security. Available at: https://www.iso.org/obp/ui/#iso:std:iso-iec:27032:ed-2:v1:en
4. Grusho, A. A., Grusho, N. A., Zabezhailo, M. I., Timonina, E. E. (2016). Intelligent data analysis in information security. Automatic Control and Computer Sciences, 50 (8), 722–725. https://doi.org/10.3103/s0146411616080307
5. Miloslavskaya, N. (2020). Stream Data Analytics for Network Attacks' Prediction. Procedia Computer Science, 169, 57–62. https://doi.org/10.1016/j.procs.2020.02.114
6. Vasilyev, V., Vulfin, A., Kuchkarova, N. (2020). Automation of Software Vulnerabilities Analysis on the Basis of Text Mining Technology. Voprosy Kiberbezopasnosti, 4 (38), 22–31. https://doi.org/10.21681/2311-3456-2020-04-22-31
7. Fatkieva, R. R., Levonevskiy, D. K. (2015). Application of Binary Trees for the IDS Events Aggregation Task. SPIIRAS Proceedings, 3 (40), 110–121. https://doi.org/10.15622/sp.40.8
8. Gonzalez Granadillo, G., El-Barbori, M., Debar, H. (2016). New Types of Alert Correlation for Security Information and Event Management Systems. 2016 8th IFIP International Conference on New Technologies, Mobility and Security (NTMS). Larnaca, 1–7. https://doi.org/10.1109/ntms.2016.7792462
9. Nevliudov, I., Yevsieiev, V., Maksymova, S., Filippenko, I. (2020). Development of an architectural-logical model to automate the management of the process of creating complex cyber-physical industrial systems. Eastern-European Journal of Enterprise Technologies, 4 (3 (106)), 44–52. https://doi.org/10.15587/1729-4061.2020.210761
10. Embracing a Zero Trust Security Model (2021). NSA. Available at: https://media.defense.gov/2021/Feb/25/2002588479/-1/-1/0/CSI_EMBRACING_ZT_SECURITY_MODEL_UOO115131-21.PDF
11. Evans, M., Maglaras, L. A., He, Y., Janicke, H. (2016). Human behaviour as an aspect of cybersecurity assurance. Security and Communication Networks, 9 (17), 4667–4679. https://doi.org/10.1002/sec.1657
12. Security and privacy controls for federal information systems and organizations (2022). U.S. Department of Commerce, Washington, D.C. NIST Special Publication 800-53, Rev 4. Available at: http://dx.doi.org/10.6028/NIST.SP.800-53r4
13. Dzheniuk, N., Yevseiev, S., Lazurenko, B., Serkov, O., Kasilov, O. (2023). A method of protecting information in cyber-physical space. Advanced Information Systems, 7 (4), 80–85. https://doi.org/10.20998/2522-9052.2023.4.11
14. Chen, Y., Zhang, Y., Wang, Z., Wei, T. (2017). Downgrade Attack on TrustZone. https://doi.org/10.48550/arXiv.1707.05082
15. Pohasii, S., Milevskyi, S., Tomashevsky, B., Voropay, N. (2022). Development of the double-contour protection concept in socio-cyberphysical systems. Advanced Information Systems, 6 (2), 57–66. https://doi.org/10.20998/2522-9052.2022.2.10
16. Zhang, M., Wang, L., Jajodia, S., Singhal, A. (2021). Network Attack Surface: Lifting the Concept of Attack Surface to the Network Level for Evaluating Networks' Resilience Against Zero-Day Attacks. IEEE Transactions on Dependable and Secure Computing, 18 (1), 310–324. https://doi.org/10.1109/tdsc.2018.2889086
17. NIST AI 100-1 Artificial Intelligence Risk Management Framework (AI RMF 1.0) (2023). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.AI.100-1
18. Zakharzhevskyy, A. G., Tolkachov, M. Yu., Dzhenyuk, N. V., Pogasii, S. S., Glukhov, S. I. (2024). The method of protecting information resources based on the semiotic model of cyberspace. Modern Information Security, 57 (1), 57–68. https://doi.org/10.31673/2409-7292.2024.010007
19. Canadian Institute for Cybersecurity (CIC) project funded by Canadian Internet Registration Authority (CIRA). Available at: https://www.unb.ca/cic/datasets/dohbrw-2020.html
20. Susto, G. A., Cenedese, A., Terzi, M. (2018). Time-Series Classification Methods: Review and Applications to Power Systems Data. Big Data Application in Power Systems, 179–220. https://doi.org/10.1016/b978-0-12-811968-6.00009-7
21. Vu, L., Pavuluri, V. N., Chang, Y., Turaga, D. S., Zhong, A., Agrawal, P. et al. (2018). A Large-Scale System for Real-Time Glucose Monitoring. 2018 48th Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshops (DSN-W), 3, 34–37. https://doi.org/10.1109/dsn-w.2018.00020
22. Yevseiev, S., Tolkachov, M., Shetty, D., Khvostenko, V., Strelnikova, A., Milevskyi, S., Golovashych, S. (2023). The concept of building security of the network with elements of the semiotic approach. ScienceRise, 1, 24–34. https://doi.org/10.21303/2313-8416.2023.002828