Model of a distributed heterogeneous system resistant to leakage of confidential information*

Oleksandr Bokhonko[1,†,*], Olha Atamaniuk[1,†], and Tomas Sochor[2,†]

[1] *Khmelnitsky National University, Khmelnitsky, Instytutska street 11, 29016, Ukraine*
[2] *European Research University, Ostrava, Czech Republic*

Abstract

Distributed heterogeneous systems play a crucial role in modern computing, enabling scalable, high-performance solutions for artificial intelligence, big data processing, and cloud-based applications. However, their inherent complexity and diverse technological components expose them to significant security risks, particularly in terms of confidential information leakage. This paper presents a novel model for a distributed heterogeneous system that enhances resistance to data breaches by leveraging a multi-agent approach. The proposed model integrates autonomous security agents responsible for real-time monitoring, anomaly detection, and dynamic access control. A mathematical framework formalizing agent interactions, decision-making strategies, and information flow is developed. Experimental validation demonstrates the system's resilience against various cyber threats, including SQL injections, malware, phishing, and brute-force attacks. The results indicate that multi-agent-based security mechanisms significantly improve threat detection accuracy and response efficiency, reducing the likelihood of unauthorized data exposure. This research contributes to the development of secure distributed computing environments by providing a scalable, adaptive, and robust architectural solution.

Keywords

distributed systems, heterogeneous computing, data security, multi-agent systems, confidential information leakage

1. Introduction

Today, distributed heterogeneous systems play a key role in various fields, as modern computing tasks require flexibility, scalability and high performance. They have become the basis for cloud computing, edge and fog computing, as well as for deploying artificial intelligence and big data processing. In modern business processes and scientific research, there is a growing need for the interaction of heterogeneous hardware and software platforms. This includes the use of CPUs, GPUs, FPGAs and other accelerators to efficiently perform resource-intensive tasks. For example, in financial technology and healthcare, machine learning models process huge amounts of data using computing resources distributed between cloud platforms and local computing nodes [1].

IT infrastructure can be considered a distributed heterogeneous system as it consists of multiple interconnected components, such as servers, networks, storage, and applications, that operate across different environments. These components often come from various vendors, use diverse technologies, and function under different protocols. The distribution aspect arises from the geographical and logical dispersion of resources, while heterogeneity is evident in the variety of hardware, software, and data formats. Effective management of such an infrastructure requires

Intelitsis'25: The 6th International Workshop on Intelligent Information Technologies & Systems of Information Security, Aprile 04, 2025, Khmelnytskyi, Ukraine

* Corresponding author.
† These authors contributed equally.

booweb24@gmail.com (O. Bokhonko); olhaatamaniuk12@gmail.com (O. Atamaniuk); tomas.sochor@eruni.org (T. Sochor)
0000-0002-7228-9195 (O. Bokhonko); 0000-0002-1704-1883 (T. Sochor)

interoperability, standardization, and robust coordination mechanisms to ensure seamless operation and security [2].

In distributed heterogeneous systems, data leakage is one of the most serious threats, as such systems combine heterogeneous hardware and software components operating in different environments. The main risks of data leakage in such systems are related to the lack of consistency in security policies, the complexity of access control, the vulnerability of individual nodes, and possible attacks on the network infrastructure [3]. Effective protection against data leakage in distributed heterogeneous systems requires a comprehensive approach that includes strict access control, end-to-end encryption, regular security monitoring, and the application of minimum privilege policies for users and services [4].

Another critical risk is insider threats and unintentional data leaks due to incorrect system configuration or human error. For example, insufficiently secured application programming interfaces (APIs) can become a point of leakage for sensitive information.

Developing a distributed heterogeneous system architecture that is resistant to confidential information leakage is a critical task in today's environment of heightened cyber threats and growing data security requirements. Such an architecture should ensure not only efficient resource management and interaction of various computing platforms, but also comprehensive protection against potential information leaks at all levels, from hardware to application [5].

One of the key reasons for the need to build such an architecture is the complexity of modern distributed systems that combine heterogeneous technological components: cloud computing, peripheral nodes, mobile devices, IoT sensors, and specialized computing accelerators (GPU, FPGA). This creates a large number of possible entry points for attackers and increases the risk of uncontrolled data leakage. An important requirement for such an architecture is the implementation of end-to-end data encryption, both during transmission between system nodes and at rest. Additionally, it is necessary to use differentiated access mechanisms based on the Zero Trust model, when no component of the system is considered trusted by default, and access is granted solely on the basis of thorough authentication and authorization. Implementing an architecture that is resistant to confidential information leakage will not only minimize the risk of data loss, but also increase user confidence, meet regulatory requirements and ensure the stability of the system in the face of dynamic threats.

The use of the Multi-Agent Systems (MAS) concept to develop the architecture of a distributed heterogeneous system resistant to confidential information leakage is a promising direction that allows increasing the level of security and adaptability of the systems of IT infrastructure.

The multiagent approach involves the use of autonomous software agents, each of which performs specific functions, has a certain level of autonomy and interacts with other agents to achieve common goals. In the context of distributed heterogeneous systems, this means that data security can be ensured through intelligent access control, monitoring of anomalous activity, and dynamic response to potential threats.

One of the main advantages of MAS is the possibility of decentralized security control, which reduces the risk of compromising the central controller and allows for more efficient threat detection and localization. For example, in distributed environments, each node or subsystem can have its own security agent that analyses traffic, checks access rights, monitors data leakage attempts, and interacts with other agents to share information about potential risks [6].

In addition, MASs can provide proactive management of sensitive information by implementing mechanisms to dynamically adjust access levels and using machine learning techniques to predict threats. For example, agents can analyze user behavioral patterns and detect anomalies that may indicate an unauthorized access attempt or data leakage.

Another important aspect is the ability of agents to learn and adapt to new threats. In this context, MAS can be integrated with artificial intelligence technologies to enhance the ability to analyze and respond to cyber threats in real time. It is also possible to use agents to manage distributed cryptographic mechanisms, such as dynamic encryption and key distribution, which provides an additional layer of protection.

In general, the concept of multi-agent systems has significant potential for developing the architecture of distributed heterogeneous systems with increased resistance to confidential information leakage. It provides autonomy, flexibility, adaptability, and decentralized decision-making, which are critical factors for data protection in complex computing environments.

2. Related works

There are a huge number of researches devoted to the leakage of confidential information problem.

In [7] a study aimed to bridge this gap by proposing a comprehensive model that examines the interrelationships between information security culture, information leakage, information sharing effectiveness, and supply chain resilience. Using a cross-sectional survey of senior managers from multinational corporations and small and medium enterprises, the researchers employed structural equation modeling to analyze the data. The findings confirmed the proposed model, demonstrating that information security culture and information leakage are negatively correlated, both of which significantly impact supply chain resilience.

A study [8] assessed various anonymization methods, such as generalization, k-anonymity, pseudonymization, and data masking, within healthcare systems. The research demonstrated that these techniques significantly mitigate the risk of data leakage while maintaining the integrity of patient information. However, the study also emphasized the challenge of maintaining system performance while ensuring robust security. The results underscored the need for a balance between data privacy and operational efficiency, showing that while anonymization enhances data protection, it can lead to performance trade-offs. These findings provide valuable insights for securing HIS without compromising the quality-of-service delivery, aligning with ongoing efforts in healthcare cybersecurity.

A study [9] introduces a novel RIS-enhanced backscatter communication system, which leverages radio frequency (RF) signals from a power beacon (PB) to securely transmit information to multiple authorized users, each using a single antenna. To optimize system performance, the study employs the Deep Deterministic Policy Gradient (DDPG) algorithm for dynamic RIS beamforming control. This approach aims to mitigate eavesdropping attempts by adversaries with linear decoding techniques. Simulation results demonstrate that the DDPG-based strategy outperforms traditional optimization methods, significantly improving multicast secrecy rates while adhering to transmit power and unit modulus constraints. The research highlights how RIS and backscatter communication can enhance security and energy efficiency in future 6G networks, offering a scalable solution to counter eavesdropping threats in emerging wireless systems.

As digital interactions continue to expand, securing data privacy and system integrity has become increasingly critical. A growing body of research has focused on advanced techniques for safeguarding digital systems. For example, studies have highlighted the role of encryption algorithms, biometric authentication, machine learning for anomaly detection, and blockchain technology in forming a robust defense against evolving cyber threats. A notable contribution [10] explores the synthesis of these techniques into a comprehensive security strategy, revealing that a holistic approach offers enhanced protection. The research emphasizes the importance of user-centric security measures, continual adaptation to emerging threats, and the ethical considerations that accompany technological advancements. By providing actionable insights, the paper offers practical recommendations for both researchers and practitioners, helping stakeholders navigate the complex landscape of data privacy and security. The study contributes to a nuanced understanding of the dynamic nature of data protection in the digital age, aiming to ensure the resilience and trustworthiness of digital systems through carefully crafted, adaptable security solutions.

A study [11] explores the factors and components that shape IT security within the framework of professional ethics policies in municipal organizations. This applied-developmental and exploratory research employs a qualitative methodology, using semi-structured interviews with experts from both academic and executive backgrounds. The research identifies key dimensions impacting IT security, including professional ethics, commitment and responsibility, creativity and innovation, human resource management, human resource performance, and organizational structure. Notably, the study found that municipalities with institutionalized professional ethics principles demonstrated a greater ability to manage and control security issues, highlighting the importance of balancing contextual, behavioral, and structural dimensions for successful IT security management. The research also reveals that organizations with stronger ethical practices are more resilient in addressing and mitigating IT security challenges, providing valuable insights for improving IT security strategies in municipal settings.

Studies [12] emphasize the difficulty of detecting data leakage in biological datasets due to their high correlation and hidden dependencies. Some works propose verification methods such as stratified cross-validation or detailed sample analysis, but there is no universal solution. In this regard, a set of seven key questions was proposed to help identify and prevent data leakage when developing machine learning models in biological applications. The practical usefulness of such approaches was demonstrated on complex examples that require in-depth analysis of the sources of leakage. The researchers note that the use of the proposed questions contributes to increasing the reliability and reproducibility of machine learning in biological research, which is critical to ensuring the reliability of the obtained results.

A study [13] addresses these security challenges by proposing a novel security scheme for protecting Video-GIS data in an open and shared environment. This scheme combines digital watermarking and data encryption to safeguard the data. Video-GIS data is categorized into general and confidential types based on the presence of sensitive information within the image frames, with tailored security measures applied to each category. Experimental results demonstrate that the watermarking algorithm has minimal impact on the quality of the data while ensuring optimal invisibility and robustness.

A study [14] explores the security issues in SIS schemes derived from AB-SS, particularly in (2, n)-CRTSIS schemes, where a vulnerability in a single share image can be exploited to reveal confidential information, including secret pixel values and their ratios. To address these security concerns, the paper proposes an enhancement to the AB-SS core sharing principle by introducing a chain obfuscation technology based on the XOR operation. The resulting secure image sharing scheme, COxor-CRTSIS, employs integer linear programming to achieve lossless recovery without segmentation and eliminates potential risks of secret disclosure without requiring additional encryption.

A study [15] proposes a solution combining cryptography and image steganography to enhance cloud data protection. This approach utilizes the Advanced Encryption Standard (AES) for encryption, ensuring that data remains unreadable to unauthorized users, while Diffie-Hellman facilitates secure key exchange to further strengthen access control. Additionally, the encrypted data is hidden within digital images using Discrete Cosine Transform (DCT) steganography, adding an extra layer of security against potential breaches. The proposed method offers an effective solution to safeguard data confidentiality, integrity, and availability without impacting system performance.

A study [16] addresses this challenge by exploring the use of deep learning for HT detection. The paper compares deep learning-based detection methods with traditional approaches and introduces the deep support vector data description (Deep SVDD) model as a novel solution. The proposed method significantly outperforms existing detection techniques, achieving an average accuracy of 92.87%, compared to 50.00% for conventional methods.

A study [17] introduces a novel training policy designed to reduce training time within an FL environment using HE, while maintaining privacy. This approach progressively reduces the amount of training data and exchanges LR coefficients in a privacy-preserving manner. The research evaluates the performance of FL policies with HE-LR, showing that the proposed policy can accelerate training times by 12% to 69% compared to traditional FL approaches, with only a slight average accuracy decrease of 1.79% to 1.95%. This contribution provides valuable insights into optimizing training efficiency while ensuring privacy in federated learning settings.

A study by [18] proposes a hybrid secure technique to protect data during NoC transmission. The proposed approach combines the Noekeon and RSA algorithms to form a hybrid security model tailored for NoC architectures. The Noekeon algorithm, known for its high security, efficiency, flexibility, and resistance to side-channel attacks, is used to secure communications within the NoC. Additionally, the RSA encryption algorithm is modified to reduce computational overhead by minimizing the number of calculations. The proposed hybrid secure algorithm is tested on a 4 × 4 2D mesh NoC architecture, showing significant improvements in performance—an increase in average throughput by 64% and a reduction in latency by 51% compared to existing methods.

A study [19] proposes a novel security system that combines multiple cryptographic algorithms and steganography to enhance data protection in cloud storage. The proposed method utilizes fast and secure symmetric key algorithms, such as AES and DES, alongside the asymmetric RSA algorithm to create a robust encryption system. This hybrid approach leverages the strengths of each algorithm to safeguard data from unauthorized access.

A study [20] addresses this challenge by proposing a System-on-Chip (SoC) architecture that integrates the SHA-256 cryptographic algorithm to enhance data security within IoT environments. The paper emphasizes the use of SHA-256 due to its strong cryptographic properties, which provide a high level of data security and integrity. The proposed architecture includes six General Purpose Input/Output (GPIO) pins, enhancing the flexibility and adaptability of IoT devices. This design also integrates a Zynq UltraScale+ MPSoC board, using SHA-256 encryption to secure sensitive data transfers through end-to-end encryption. The system is further optimized with a Verilog implementation of the SHA-256 block, employing GPIO for input and I2C for communication with a camera, while utilizing SRAM connected to registers and an ALU. UART is used for output transfer, enabling further processing and analysis. The results demonstrate that the architecture not only offers robust security but also provides excellent power efficiency and performance, making it well-suited for a wide range of IoT applications.

A study [21] proposes a robust ISS designed specifically for cloud computing, focusing on the integration of cryptographic techniques. The research combines the RSA-OAEP (Optimal Asymmetric Encryption Padding) algorithm with X.509 to develop a comprehensive security system. The paper further explores the role of various cryptographic methods, such as encryption, digital signatures, and key management, in protecting cloud-based systems. In addition, it discusses enhancing the detection and response capabilities of security systems by incorporating artificial intelligence (AI) algorithms, specifically Grubbs' Test, to identify potential threats.

A study [22] introduces the OctagonCryptoDataMR paradigm, which integrates cryptographic hash and encryption/decryption techniques within the MapReduce framework to enhance data security. The proposed model uses a position-based swing hill cipher at the "diminish" layer and a sequence rolling twofold hash technique at the "map" layer to safeguard data privacy. This approach employs straightforward cryptographic techniques to achieve complex and effective data security outcomes. Experimental results show that the proposed system not only ensures data privacy but also improves processing speed with minimal execution time, making it an efficient solution for enhancing data security in cloud computing environments.

The reviewed studies highlight the growing importance of securing digital systems, particularly in distributed environments where information security threats, such as data leakage, eavesdropping, and unauthorized access, continue to evolve. Research demonstrates that existing techniques, including anonymization, cryptographic methods, and artificial intelligence-driven security measures, contribute to enhancing confidentiality and resilience. However, challenges remain, particularly in balancing security, performance, and usability. Given the increasing complexity of IT infrastructures – characterized by their distributed and heterogeneous nature – there is a critical need to develop new approaches for modeling such systems to ensure resilience against confidential information leakage. A robust model should incorporate dynamic security strategies, adaptive encryption techniques, and intelligent threat detection mechanisms while maintaining system performance. Furthermore, integrating emerging technologies such as federated learning, blockchain, and privacy-preserving computation can enhance security without compromising efficiency.

Thus, research should focus on designing comprehensive frameworks that address the unique challenges posed by distributed heterogeneous systems, ensuring confidentiality, integrity, and availability in increasingly interconnected digital environments.

3. Model of a distributed heterogeneous system resistant to leakage of confidential information

3.1. Mathematical model of IT infrastructure as a distributed heterogeneous system a multi-agent system

Let us present the IT infrastructure in terms of a distributed heterogeneous system as a multi-agent system [23]. IT infrastructure of an enterprise includes several components, each of which performs a specific function in the system. A multi-agent system (MAS) provides distributed data processing, process automation, and adaptability to changes.

Let us present the mathematical model of a multi-agent system (MAS) is based on the formalization of the interaction between agents, the description of their behavior, goals, and environment. Here are the main components of the mathematical model of such a system:

$$\mathcal{A} = \{A_1^u, A_2^s, A_3^m, A_4^c\}, \tag{1}$$

where $\mathcal{A}$ is a set of agents, that include:
A_1^u – set of *user agents* that interact with users to perform tasks such as processing requests, providing information, or supporting decisions;
A_2^s – set of *service agents*, that provide access to external and internal services;
A_3^m – set of *monitoring agents*, that monitor the status of the system, network, or individual components, analyze log files, and warn about possible failures;
A_4^c – set of *communication agents*, that are responsible for exchanging data between different agents using communication protocols.

Each agent A_i is defined as:

$$A_i = \langle S_i, E_i, \Pi_i, \Phi_i, \Omega_i, \mathrm{C}_i \rangle, \tag{2}$$

where:
S_i – set of agent states A_i;
E_i – the set of actions that an agent can perform;
Π_i – the agent's strategy or policy that determines the choice of action in a particular state:

$$\Pi_i : S_i \times \mathcal{O} \rightarrow E_i , \tag{3}$$

where $\mathcal{O}$ – agent observation;
Φ_i – utility function or objective function that defines the agent's goals;
Ω_i – a set of resources available to the agent.
C_i – a communication model that defines how an agent exchanges information with other agents.

Let us describe the environment with its states as a set:

$$\mathrm{S} = \{S_1, S_2, \dots\ S_m\}. \tag{4}$$

Let us describe the environmental dynamics function as:

$$T : S \times A \rightarrow S, \tag{5}$$

where $T$is the change in the state of the environment as a result of the actions of agents.

The interaction between agents can be presented as the communication graph:

$$\mathcal{G} = (\mathcal{A}, \mathcal{E}), \tag{6}$$

where $\mathcal{E}$is the set of edges representing the connections between agents. *An* edge e_{ii} $\in \mathcal{E}$ means that the agent A_i can be exchange via information A_j.

In order to describe the system dynamics, let us present the agent state transition model.

The state of each agent changes according to the function:

$$S_i^{t+1} = f(S_i^t, E_i^t, O_i^t, C_i^t), \tag{7}$$

where:

S_i^t – the state of the agent at time t;

E_i^t – the action chosen by the agent at the moment t;

O_i^tan agent's observations of the environment or other agents;

C_i^t – information received through communication.

To describe the interaction of an agent with the environment let us present the result of the agent's actions on the environment as:

$$S_i^{t+1} = T \ (S^t\{E_1^t, E_2^t, \dots \ E_n^t\}). \tag{8}$$

3.2. Goals and optimization

To describe the efficient system functioning let us present the *utility function*, where the agent's A_i goal is to maximize it:

$$U_i = \sum_{t=0}^{\infty} \gamma^t \ \Phi_i \ (S_i^t, S^t) \ , \tag{8}$$

where:

$\Phi_i \ (S_i^t, S^t)$ – the agent's utility at time t, γ ∈ [0,1] is the discount factor (for long-term or short-term planning).

Collective utility function If the system is focused on collective goal achievement, a global utility function is introduced:

$$U_{global} = \sum_{i=1}^{n} U_i. \tag{9}$$

Concerning the optimization, **t**he task of a multi-agent system is to find a set of strategies $\{\Pi_1, \Pi_2, \dots, \Pi_n,\}$ that maximizes the utility function:

$$arg \underset{\{\Pi_1, \dots \ \Pi_n\}}{max} U_{global}, \tag{10}$$

3.3. Distributed model

To present the distributed data processing let us show how each agent performs local optimization:

$$\text{arg} \underset{\Pi_i}{max} U_i \ . \tag{11}$$

Provided that there is consistency with the global goal through a mechanism of communication and joint information exchange.

To ensure balance in the system, the interaction of agents can be described through a Nash equilibrium, when no agent can improve its outcome by changing only its own strategy:

$$U_i \left(\Pi_i^*, \Pi_{\_i}^*\right) \ \geq \ U_i \left(\Pi_i^*, \Pi_{\_i}^*\right), \tag{12}$$

WhereΠ$_{\_i}$ are the strategies of all agents except A_i.

3.4. Agent communication and consistency

To achieve the goals, a consensus algorithm for the exchange protocols is used:

$$x_i^{t+1} = \sum_{j \in N_i} w_{ij}\ x_i^{t+1}, \tag{13}$$

where $N_i$are the agent's neighbors A_i, w_{ij} are the weights of the communication graph.

4. Monitoring and ensuring data integrity

To improve security and control over the enterprise's IT infrastructure, the multi-agent system is supplemented with monitoring and data integrity functionality. The main goal is to detect anomalies, prevent data theft, and ensure the system's resilience to cyberattacks.

Let us denote the agents and their functions for monitoring via the *Monitoring agent* A_{mon}, which monitors system logs, network traffic, and user activity, uses anomaly detection techniques (machine learning, threshold models), and responds to suspicious activity by generating alarms; the *Integrity agent* A_{int}, which uses hash functions $H(x)$ to verify the integrity of critical data, compares checksums of files and databases in real time, triggers a recovery mechanism when changes are detected in critical files; and the *Security agent* A_{sec}, which controls access to confidential information using authentication and authorization policies, blocks suspicious connections and requests that may cause data leakage, uses cryptographic algorithms to encrypt data transmitted between agents.

The multi-agent system is supplemented with new components, and its model is expanded with the anomaly monitoring, where each user request is modeled as a vector of behavioral parameters:

$$x = \{T_{access}, R_{data}, P_{usage}\}, \tag{14}$$

where T_{access} is the access time, R_{data} is the data volume, P_{usage} is the type of request.

Let us define the anomaly detection function as follows:

$$A(x) = \begin{cases} 1, if\ d(x,\mu) < \lambda \\ 0, \text{otherwise} \end{cases}, \tag{15}$$

where $d(x, \mu)$ is the distance from the average behavioral profile, λ is the anomaly threshold.

Let us present the data integrity verification model. For each critical file or database record, a hash sum is calculated:

$$H(D) = SHA256(D), \tag{16}$$

and the comparison of checksums at different points in time:

$$\Delta H = H(Dt) - H(Dt - 1), \tag{17}$$

Where if ΔH≠0Δ, verification is started and a rollback to the backup copy is possible.

Access is controlled through the RBAC (Role-Based Access Control) policy:

$$P(u, r) = \begin{cases} 1, if\ u \in R_{allowed} \\ 0, \text{otherwise} \end{cases}, \tag{18}$$

where u is the user, $R_{allowed}$ is the set of allowed roles for accessing the resource.

5. Experiments

5.1. MAS stability test

Experimental studies of the proposed multi-agent model of the enterprise IT infrastructure involved testing its efficiency, stability and adaptability in various operating scenarios. For this purpose, a simulation environment (Matlab/Simulink [24, 25]) was deployed that simulated the real infrastructure, including databases, network interaction and communication between agents. For this purpose, a set of test scenarios was created that take into account normal and critical operating conditions.

At the initial stage, agents were initialized, which interact with each other and with the environment according to the defined rules. The next step was to determine performance metrics such as query processing time, resource usage and fault tolerance level.

The first experiment was aimed at assessing the basic performance of the system at nominal load. Each agent performed its functions under standard conditions, and key performance indicators were recorded.

Next, a series of stability experiments were conducted, simulating failures of individual system components, server failures, or loss of communication between agents. The model was evaluated by the speed of recovery and the efficiency of load balancing between agents. Special attention is paid to self-healing mechanisms and automatic task redistribution. Another direction was the study of the scalability of the system.

In this experiment, the number of agents and the volume of processed requests gradually increased. At the same time, communication delays, the efficiency of computing resources, and stability of operation with increasing load were analyzed.

The last stage included testing the system's adaptability to changes in the environment. For this, changes were made to the rules of interaction between agents, the configuration of resources, and network parameters, which allowed us to assess the system's ability to adapt to new conditions. The results obtained are compared with the expected indicators, and on this basis, conclusions are drawn about the effectiveness of the proposed model.

The following key metrics are used to evaluate the performance of the proposed multi-agent system of the enterprise IT infrastructure:

- request processing time T_{resp} - the average time required by the agent to execute the received request;
- system throughput R_{sys} - the number of requests processed by the system per unit of time;
- agent load U_A - the average level of resource utilization of each agent;
- Failure tolerance S_{fail} - the probability of correct system operation when some of the agents fail;
- scalability S_{scal} - the dependence of performance on the number of agents and load;
- recovery time T_{rec} - the average time required to restore the system after a failure;
- Communication efficiency $Comm_{eff}$ - the average data transfer time between agents.

MAS stability test results are presented in Table 1.

Table 1
MAS stability test results

Experiment	T_{resp}, sec	R_{sys}, requests/sec	U_A, %	S_{fail}, %	S_{scal}	T_{rec}, sec	$Comm_{eff}$, ms
Rated load	0.35	200	65	99.8	1.0	0.0	5.2
High Load (×2)	0.45	380	80	99.5	0.95	0.0	6.5

10% Agent Opt-Out	0.40	190	70	98.0	0.90	1.2	5.8
30% Rejection of Agents	0.55	150	75	92.5	0.85	2.8	6.9
Scaling (+50% agents)	0.38	300	60	99.7	1.2	0.0	4.8
Changing the environment	0.42	210	68	99.0	0.98	0.5	5.5

The system demonstrates high efficiency at nominal load, quickly adapts to changes and has a high level of resilience to agent failures. The scalability of the system allows for increased throughput when adding new agents, although communication efficiency decreases somewhat with a significant increase in load. Recovery time after failures remains within acceptable limits, which confirms the reliability of the multi-agent model.

5.2. IT-infrastructure data leak attack test

Let us describe the scenarios for executing an experiment when attacking IT infrastructure. To evaluate the operation of a multi-agent system with data monitoring and protection functions, a series of experiments is carried out aimed at determining the time of attack detection, detection accuracy, speed of data integrity verification and success of recovery after an attack.

5.2.1. Scenario 1. SQL injection data leak attack

An attacker carries out an SQL injection attack in an attempt to gain unauthorized access to sensitive information by inserting malicious SQL queries into form input fields. The course of the experiment:

1. The monitoring agent A_{mon} analyzes database queries.
2. Abnormal behavior is detected using a threshold model of deviations.
3. Security agent A_{sec} blocks a suspicious request.
4. The integrity agent A_{int} checks if there have been any changes to the database.

Test results are presented in Table 2.

Table 2
Test results for SQL injection data leak attack

No	Anomaly detection time (s)	Detection accuracy (%)	Integrity check time (ms)	Recovery success rate (%)
1	0.32	97.8	4.8	100.0
2	0.34	97.9	4.9	99.8
3	0.29	98.1	4.6	100.0
4	0.33	97.5	4.8	99.9
5	0.30	98.0	4.5	100.0
6	0.32	97.7	4.8	99.9
7	0.35	97.8	5.0	99.7
8	0.28	98.2	4.4	100.0
9	0.33	97.9	4.9	99.8
10	0.31	97.6	4.7	100.0

5.2.2. Scenario 2. Data leak due to malware

A malicious script is downloaded to the server that secretly copies sensitive files and sends them to an external server.

The course of the experiment:

1. The monitoring agent A_{mon} analyzes files sent outside the network.
2. Abnormal use of resources and network traffic is detected.
3. The security agent A_{sec} isolates the process and terminates the connection.
4. The integrity agent A_{int} checks the hashes of the files and performs the recovery.

Test results are presented in Table 3.

Table 3
Test results for data leak due to malware

No	Anomaly detection time (s)	Detection accuracy (%)	Integrity check time (ms)	Recovery success rate (%)
1	0.32	97.8	4.8	100.0
2	0.43	98.6	5.7	99.3
3	0.44	98.7	5.4	99.2
4	0.42	98.3	5.2	99.1
5	0.47	98.5	5.6	99.4
6	0.45	98.2	5.5	99.2
7	0.48	98.8	5.8	99.5
8	0.41	98.3	5.1	99.0
9	0.49	98.7	5.9	99.6
10	0.44	98.6	5.4	99.3

5.2.3. Scenario 3. Data theft due to a phishing attack

The attacker sends an email with a fake link that directs the employee to a malicious site to enter credentials.

The course of the experiment:

1. The monitoring agent A_{mon} scans emails for suspicious attachments and links.
2. A URL leading to a phishing site is revealed.
3. The A_{sec} security agent blocks access to the site.
4. The integrity of credentials and the user's login history are checked.

Test results are presented in Table 4.

5.2.4. Scenario 4: Using stolen credentials (Brute-Force Attack)

The attacker tries to guess the administrator's password by repeatedly trying to log in (brute-force).

The course of the experiment:

1. The monitoring agent A_{mon} captures a suspicious number of failed login attempts.
2. The security agent A_{sec} automatically blocks the attacker's IP address.
3. The integrity agent A_{int} checks to see if there have been any changes to the access entries.

Test results are presented in Table 5.

Table 4

Test results for data theft due to a phishing attack

No	Anomaly detection time (s)	Detection accuracy (%)	Integrity check time (ms)	Recovery success rate (%)
1	0.39	96.7	4.2	98.5
2	0.38	96.9	4.3	98.6
3	0.37	97.0	4.2	98.4
4	0.39	96.8	4.1	98.5
5	0.41	96.6	4.2	98.3
6	0.36	97.1	4.0	98.7
7	0.42	96.4	4.4	98.2
8	0.37	97.0	4.1	98.6
9	0.40	96.7	4.3	98.5
10	0.38	96.9	4.2	98.6

Table 5

Test results for using stolen credentials

No	Anomaly detection time (s)	Detection accuracy (%)	Integrity check time (ms)	Recovery success rate (%)
1	0.27	99.0	5.0	99.7
2	0.29	99.2	5.2	99.9
3	0.30	99.3	5.3	99.8
4	0.28	99.1	5.1	99.8
5	0.26	99.4	4.9	99.7
6	0.31	99.2	5.4	99.9
7	0.29	99.0	5.3	99.8
8	0.27	99.3	5.0	99.9
9	0.32	99.1	5.5	99.7
10	0.28	99.4	5.1	99.8

The system effectively detects various types of attacks, ensuring a quick response and minimizing the risk of data theft. The best results are demonstrated in brute-force attacks and SQL injections thanks to operational monitoring of requests. The longest detection time was recorded in an attack through malicious software, which is associated with the need to analyze network traffic. In general, the system is able to respond quickly to threats, block abnormal activity and guarantee a high success rate of data recovery in the event of an attack.

Conclusion

This study introduced a novel model for a distributed heterogeneous system designed to enhance cybersecurity through a multi-agent approach. By incorporating autonomous security agents, the system enables real-time monitoring, anomaly detection, and adaptive access control, significantly improving resilience against cyber threats such as SQL injections, malware, phishing, and brute-force attacks. The mathematical framework developed in this research formalizes agent interactions and decision-making, ensuring an adaptive and scalable security mechanism.

Experimental validation confirms that the proposed approach enhances threat detection accuracy and reduces response time, minimizing the risk of unauthorized data exposure. The findings of this research contribute to the development of more secure distributed computing environments, offering a robust and flexible security architecture suitable for modern digital infrastructures. Future work may explore the integration of machine learning techniques to further optimize agent-based security mechanisms.

Declaration on Generative AI

During the preparation of this work, the authors used Grammarly in order to: grammar and spelling check; DeepL Translate in order to: some phrases translation into English. After using these tools/services, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] O. Revniuk, A. Postoliuk, Research on the application of adaptive risk assessment methods for web applications, Computer Systems and Information Technologies, 2024 (3), 34–43. https://doi.org/10.31891/csit-2024-3-5.

[2] O. Savenko, A. Sachenko, S. Lysenko, G. Markowsky, N. Vasylkiv, Botnet detection approach based on the distributed systems, International Journal of Computing, 2020 (2), 190–198. https://doi.org/10.47839/ijc.19.2.1761.

[3] O. Savenko, S. Lysenko, A. Kryschuk, Multi-agent based approach of botnet detection in computer systems. In: Communications in Computer and Information Science 291 (2012) 171–180. https://doi.org/10.1007/978-3-642-31217-5_19.

[4] S. Lysenko, O. Savenko, K. Bobrovnikova, DDoS botnet detection technique based on the use of the semi-supervised fuzzy c-means clustering, CEUR-WS, 2018 (2104), 688–695.

[5] S. Lysenko, O. Savenko, K. Bobrovnikova, A. Kryshchuk, B. Savenko, Information technology for botnets detection based on their behaviour in the corporate area network, Communications in Computer and Information Science, 2017 (718), 166–181.

[6] Z. Xia, W. Yu, Y. Liu, J. Lü, Distributed bilevel constrained optimization via multiagent system approaches, IEEE Transactions on Cybernetics, 2025. https://doi.org/10.1109/TCYB.2025.3531393.

[7] W. P. Wong, K. H. Tan, K. Govindan, et al., A conceptual framework for information-leakage-resilience, Annals of Operations Research, 329 (2023), 931–951. https://doi.org/10.1007/s10479-021-04219-5.

[8] Rapšík, M. Kvet, Improving cybersecurity in hospital information systems through anonymization techniques, 2025 IEEE 23rd World Symposium on Applied Machine Intelligence and Informatics (SAMI), Stará Lesná, Slovakia, 2025, 447–452. https://doi.org/10.1109/SAMI63904.2025.10883123.

[9] S. Z. U. A. Abideen, A. Wahid, M. M. Kamal, et al., Advancements in IoT system security: a reconfigurable intelligent surfaces and backscatter communication approach, Journal of Supercomputing, 81 (2025), 362. https://doi.org/10.1007/s11227-024-06819-x.

[10] R. Kumari, S. Sriramulu, Exploring advanced techniques for enhancing data privacy and security in digital systems, 2024 1st International Conference on Advances in Computing, Communication and Networking (ICAC2N), Greater Noida, India, 2024, 327–334. https://doi.org/10.1109/ICAC2N63387.2024.10895522.

[11] X. Wang, T. Ahonen, J. Nurmi, Applying CDMA technique to network-on-chip, IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 15(10), 2007, 1091–1100.

[12] J. Bernett, D. B. Blumenthal, D. G. Grimm, et al., Guiding questions to avoid data leakage in biological machine learning applications, Nature Methods, 21(8), 2024, 1444–1453.

[13] O. Mykhaylova, M. Korol, R. Kyrychok, Research and analysis of issues and challenges in ensuring cybersecurity in cloud computing, Cybersecurity Providing in Information and Telecommunication Systems II, 3826, 2024, 30–39.

[14] Y. Qiu, J. Long, C. Ma, et al., Security protection of video-GIS data based on data encryption and digital watermarking, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLVIII-4-2024, 681–688. https://doi.org/10.5194/isprs-archives-XLVIII-4-2024-681-2024.

[15] R. Wang, L. Li, G. Yang, et al., IEEE Transactions on Information Forensics and Security, 19, 2024. https://doi.org/10.1109/TIFS.2024.3477265.

[16] G. Shidaganti, M. V. L., M. Vinay, P. Patil, 2024 5th International Conference on Circuits, Control, Communication and Computing (I4C), Bangalore, India, 2024. https://doi.org/10.1109/I4C62240.2024.10748507.

[17] D. Lee, J. Lee, Y. Jung, J. Kauh, T. Song, IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 32(12), 2024. https://doi.org/10.1109/TVLSI.2024.3458892.

[18] J. M. Cortés-Mendoza, A. Tchernykh, H. González-Vélez, 2024 2nd International Conference on Federated Learning Technologies and Applications (FLTA), Valencia, Spain, 2024. https://doi.org/10.1109/FLTA63145.2024.10839854.

[19] T. Nagalaxmi, E. S. Rao, P. ChandraSekhar, FPGA-based implementation and verification of hybrid security algorithm for NoC architecture, Volume 121, 2024, 13–23.

[20] P. Neelakantan, G. Malige, 2024 International Conference on Emerging Techniques in Computational Intelligence (ICETCI), Hyderabad, India, 2024. https://doi.org/10.1109/ICETCI62771.2024.10704081.

[21] D. Ferlin Deva Shahila, L. Padma Suresh, P. Aruna Jeyanthy, V. Stephen, A. R. M., 2024 7th International Conference on Circuit Power and Computing Technologies (ICCPCT), Kollam, India, 2024. https://doi.org/10.1109/ICCPCT61902.2024.10673314.

[22] V. Srilakshmi, V. L. Chetana, R. R. Dornala, S. P. Vallabaneni, G. Ramanjaiah, 2024 5th International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2024. https://doi.org/10.1109/ICESC60852.2024.10690012.

[23] G. S. Brindha, G. Siva, Octagon-CryptoDataMR: Enhanced data protection using octagon-based cryptographic hash with encryption and decryption through map reduce programming model, Volume 56(1), May 2024.

[24] W. He, W. Xu, X. Ge, et al., Secure control of multiagent systems against malicious attacks: A brief survey, IEEE Transactions on Industrial Informatics, 18(6), 2022, 3595–3608. https://doi.org/10.1109/TII.2021.3126644.

[25] V. Saini, P. Shah, R. Sekhar, MATLAB and Simulink for building automation, 2022 IEEE Bombay Section Signature Conference (IBSSC), Mumbai, India, 2022, 1–6. https://doi.org/10.1109/IBSSC56953.2022.10037485.

[26] X. Ban, M. Ding, S. Liu, et al., TAESim: A testbed for IoT security analysis of trigger-action environment, European Symposium on Research in Computer Security, Cham: Springer International Publishing, October 2021, 218–237.