Information systems. Computer sciences. Issues of information security
Информационные системы. Информатика. Проблемы информационной безопасности

UDC 004.056.53
https://doi.org/10.32362/2500-316X-2026-14-1-19-30
EDN TXHMHW

RESEARCH ARTICLE

Identification of the message flow between two subscribers in multi-agent systems based on the analysis of its contextual characteristics

Maxim O. Tanygin,
Ilya O. Mishin [@],
Elena A. Kuleshova,
Alexey V. Kiselev

Southwest State University, Kursk, 305040 Russia
[@] Corresponding author, e-mail: mishin.ilya46@yandex.ru

• **Submitted:** 28.04.2025 • **Revised:** 01.07.2025 • **Accepted:** 17.11.2025

Abstract

Objectives. The paper examines the problem of improving the accuracy of identifying the message flow between two subscribers in multi-agent systems. This is done by analyzing the contextual characteristics of the overall message flow in the communication channel. Situations may arise during the process of source identification and verification of authenticity in which authentication codes for two or more messages collide. One way to resolve such conflicts is to isolate the message flow between two subscribers by leveraging its unique statistical characteristics which differ from the characteristics of the general message flow within the system. The aim of the paper is to develop a method which reliably identifies the target message flow even in cases of authentication code collisions.

Methods. The contextual characteristics of messages are used, in order to analyze and highlight patterns of agent behavior in the message flow. These characteristics include frequency of sending, message size, timestamps, and historical interaction data. The method involves the formation of statistical characteristics of the message flow between two agents in a multi-agent system, such as skewness and kurtosis, as well as distribution parameters for the number of messages sent between events from the target source along with their classification by means of logistic regression.

Results. During experiments, the method developed has been found to demonstrate a Precision metric value in the range of 0.81–0.85. This is 40–50% higher than existing methods based on the analysis of inter-packet time intervals, indicating that 81–85% of the messages classified as belonging to the target source are actually such. ROC[1] analysis confirmed the high efficiency of the model and acceptable classification quality.

Conclusions. The results of the study show that the use of contextual characteristics and statistical analysis enables the accurate identification of target flows in multi-agent systems with a total number of agents ranging from 70 to 110. This method can be used in low-bandwidth communication channels where it is essential to minimize the size of the transmitted batch header and computational costs associated with authentication procedures.

[1] Receiver operating characteristic.

Keywords: multi-agent systems, contextual characteristics, binary classification, logistic regression, skewness, kurtosis, ROC analysis

For citation: Tanygin M.O., Mishin I.O., Kuleshova E.A., Kiselev A.V. Identification of the message flow between two subscribers in multi-agent systems based on the analysis of its contextual characteristics. *Russian Technological Journal.* 2026;14(1):19–30. https://doi.org/10.32362/2500-316X-2026-14-1-19-30, https://www.elibrary.ru/TXHMHW

Financial disclosure: The authors have no financial or proprietary interest in any material or method mentioned.

The authors declare no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

Выделение потока сообщений между двумя абонентами в многоагентных системах на основе анализа его контекстуальных характеристик

М.О. Таныгин,
И.О. Мишин [@],
Е.А. Кулешова,
А.В. Киселев

Юго-Западный государственный университет, Курск, 305040 Россия
[@] Автор для переписки, e-mail: mishin.ilya46@yandex.ru

• **Поступила:** 28.04.2025 • **Доработана:** 01.07.2025 • **Принята к опубликованию:** 17.11.2025

Резюме

Цели. В статье исследуется задача повышения точности выделения потока сообщений между двумя абонентами в многоагентных системах на основе анализа контекстуальных характеристик общего потока сообщений в канале связи. При проведении процедур определения источника и установления его подлинности могут возникать коллизии кодов аутентификации двух и более сообщений. Одним из способов разрешения подобных ситуаций является выделение потока сообщений между двумя абонентами на основе его статистических характеристик, отличающихся от характеристик общего потока сообщений в системе. Цель работы – разработка метода, позволяющего надежно идентифицировать целевой поток в случае возникновения коллизий кодов аутентификации.

Методы. Для анализа и выделения паттернов активности агентов в потоке сообщений использованы контекстуальные характеристики сообщений: частота отправки, размер, временные метки и исторические данные взаимодействий. Метод включает формирование статистических характеристик потока сообщений между двумя агентами многоагентной системы (коэффициенты асимметрии и эксцесса, параметры распределения количества сообщений между событиями целевого источника) и их классификацию с помощью логистической регрессии.

Результаты. В ходе проведенных экспериментов было установлено, что разработанный метод демонстрирует значения метрики Precision (полнота) в диапазоне 0.81–0.85 (от 81% до 85% сообщений, классифицированных как принадлежащие целевому источнику, действительно являются таковыми), что на 40–50% превышает показатели существующих методов, основанных на анализе межпакетных интервалов времени. ROC[2]-анализ подтвердил высокую эффективность модели и приемлемое качество классификации.

[2] Receiver operating characteristic – рабочая характеристика приемника.

Выводы. Результаты исследования показали, что использование контекстуальных характеристик и статистического анализа позволяет точно выделять целевые потоки при общем числе агентов в многоагентных системах от 70 до 110. Метод может применяться в каналах связи с низкой пропускной способностью, где необходимо минимизировать размер заголовочных частей передаваемых пакетов данных и вычислительные затраты на выполнение процедур аутентификации.

Ключевые слова: многоагентные системы, контекстуальные характеристики, бинарная классификация, логистическая регрессия, асимметрия, эксцесс, ROC-анализ

Для цитирования: Таныгин М.О., Мишин И.О., Кулешова Е.А., Киселев А.В. Выделение потока сообщений между двумя абонентами в многоагентных системах на основе анализа его контекстуальных характеристик. *Russian Technological Journal*. 2026;14(1):19–30. https://doi.org/10.32362/2500-316X-2026-14-1-19-30, https://www.elibrary.ru/TXHMHW

Прозрачность финансовой деятельности: Авторы не имеют финансовой заинтересованности в представленных материалах или методах.

Авторы заявляют об отсутствии конфликта интересов.

INTRODUCTION

A multi-agent system (MAS), often referred to as a self-organized system, represents an advanced distributed computing network composed of numerous interacting agents. Each agent operates with a degree of autonomy and the capacity to adjust to its surroundings. MAS is used in a variety of fields including robotics, resource management, intelligent transportation, and e-commerce. The effectiveness of MAS systems relies significantly on the reliability and safety of interactions between agents, thereby making authentication a vital factor to consider [1].

In MAS, authentication focuses on verifying the identity of data sources, known as network agents. This process is essential for safeguarding the trustworthiness and privacy of data during transmission. With the escalation of cybersecurity threats, the development of adaptable and resilient authentication methods is increasingly crucial.

In MAS, multiple authentication techniques are employed, each with its own influence on the security and operational efficiency of the system.

Cryptography is a common approach to authentication in MAS. The use of both symmetric and asymmetric cryptographic algorithms enables trustworthy data security and agent identification [2]. Symmetric approaches, exemplified by the advanced encryption standard (AES), deliver rapid processing speeds, although they require a secure key exchange. Asymmetric encryption techniques such as Rivest–Shamir–Adleman (RSA) and elliptic-curve cryptography circumvents the challenges of key distribution. In resource-limited environments, they might prove less efficient [3].

A variety of authentication protocols have been specifically designed for MAS. For example, challenge-response protocols enable agents to verify their identity while keeping their secret keys private. These protocols may be enhanced with timestamp mechanisms to defend against replay attacks [4, 5].

Multifactor authentication (MFA), also employed in MAS systems, demands confirmation through multiple factors, in order to validate the identity of an agent. While MFA significantly enhances security within MAS, implementation introduces complexity to the authentication procedure and potentially increases system response times. It is important to acknowledge that design or implementation shortcomings in the authentication protocols can result in vulnerabilities and performance problems. Study [6] explores the challenge of identifying weaknesses in MFA protocols by means of a structured analytical approach. The authors establish a comprehensive set of criteria to evaluate security, encompassing both established and novel factors. The paper further examines the applicability of MFA across diverse sectors and pinpoints potential vulnerabilities. The study also uncovers critical flaws in ten prominent MFA protocols and offers solutions to mitigate these risks.

Dynamic authentication is another noteworthy aspect of MAS which uses fluctuating parameters to verify agent identities. This can include temporary passwords or unique codes generated through specific algorithms. This method significantly reduces the vulnerabilities arising from compromised static credentials. For example, in article [7] the author examines passwords as a form of authentication. Despite advancements in authentication methods, this technique remains the primary authentication approach in numerous systems. When heightened security is necessary, it is often used together with other authentication methods, creating a two-factor or multi-factor authentication setup. This research highlights the drawbacks of the current authentication method and suggests an alternative of a time-based one-time password (TOTP) system. TOTP generates unique access codes which are only valid for a specific timeframe. This approach, along with dynamic passwords and a hybrid system blending one-time passwords with traditional passwords, offers solutions for both user authentication and peer-to-peer (P2P) authentication.

In MAS, where agents interact with each other under uncertain conditions, reputation-based authentication plays a crucial role. This process enables agents to evaluate the reliability and trustworthiness of other agents based on their past interactions. The approach not only allows for the authentication of agents but also fosters the development of trust relationships within the system. Study [8] proposes a method for message authentication using blockchain technology. This method relies on a reputation assessment mechanism, in order to authenticate messages. Blockchain is employed to manage identities and certificates in this context, while identity authentication and certificate revocation are facilitated through smart contracts. A reputation mechanism is devised in the aims of evaluating the credibility of messages. This can be integrated with message authentication, in order to ensure effective verification of message reliability. The performance analysis conducted in the study demonstrates that the proposed approach is more efficient than the traditional signature-based authentication method. This has been determined by comparing the following characteristics: computational power requirements, message latency, and data loss rate. The findings of the authors indicate that the solution proposed offers an acceptable level of protection against various well-known types of cyberattacks.

With the advancement of blockchain technology, it is now possible to use decentralized authentication methods in MAS. Blockchain provides for both transparency and the immutability of records, enabling agents to exchange sensitive information securely. This approach has the potential to significantly enhance security and reduce risks associated with centralized management. Study [9] proposes a novel (MFA) strategy to enhance security in dynamic environments, utilizing blockchain technology. The research also introduces a consensus model based on the Raft algorithm for selecting a trusted host, ensuring fast and reliable authentication. The analysis demonstrates that this approach effectively mitigates cyberattacks which target authentication systems.

The analysis also indicates that selecting an authentication method for MAS largely hinges on the particular needs for security, performance, and scalability. Integrating multiple approaches can result in more resilient and secure systems, capable of operating efficiently in dynamic environments prone to potential threats.

OVERVIEW OF AUTHENTICATION COLLISION PROTECTION METHODS

Each message in MAS can be assigned a unique identifier, enabling it to be distinguished from other messages and preventing duplication. However, one issue which arises during the processes of source verification and authenticity validation in MAS is the occurrence of identifier collisions. For the purpose of this discussion, a collision refers to the matching of test sequences for two or more distinct messages [10]. Numerous methods exist to safeguard against such collisions. Let us examine the primary approaches.

Using unique message identifiers as verification sequences helps to prevent collisions during the authentication process. These identifiers can be generated using time stamps, random numbers, or a combination of both [11].

One effective method for guarding against collisions is the implementation of cryptographic hash functions with robust properties to mitigate the likelihood of generating identical hash outputs for distinct messages [12]. Prominent examples of such hash functions include SHA-256 and SHA-3 (Secure Hash Algorithm) [13]. The same category of applications also incorporates digital signature algorithms across various domains, in order to establish links between these algorithms and their scenarios [14]. The authors have conducted comparative analyses of widely-used digital signature algorithms such as RSA, Lamport, Elliptic Curve Digital Signature Algorithm (ECDSA), and Edwards-Curve Digital Signature Algorithm, in terms of their performance. Study [15] also analyzes existing encryption techniques and proposes a multi-layered encryption strategy. This strategy uses AES for data encryption, RSA for key protection, and role-based encryption for access control. This hierarchical method significantly raises the computational difficulty of unauthorized access attempts. Integrating these technologies has led to a 50% reduction in key compromise risks and an improvement in data integrity by leveraging MAS verification.

At the same time, when considering an MAS class where communication occurs over low-bandwidth channels, such as sensor networks, industrial Internet systems, or the Internet of Things, the algorithms mentioned above are not applicable. This limitation arises due to the restricted size of the transmitted batch header. In such cases, the likelihood of collisions is influenced more by the identifier size than the cryptographic function properties. This mechanism offers protection against spoofing and collisions, since altering the message results in a hash mismatch [16]. The paper addresses both the confidentiality of message content and the role of digital signatures in verifying its authenticity and integrity. These mechanisms not only shield the message from unauthorized access but also confirm that it has been sent by the specific agent associated with the signature.

Enhancing the reliability of flow separation—identifying the source of each message in a communication

flow—can be achieved by incorporating the contextual characteristics of the messages into the authentication process. Contextual details such as message timing, agent location, message type, and content contribute to more precise assessments of risks and subsequent actions during authentication. For example, the research outlined in [17] addresses inter-operability issues in MAS related to communication, coordination, and adaptability in dynamic environments. The study emphasizes strategies designed to elevate communication within MAS by focusing on protocols, reward structures, learning algorithms, and trust-building mechanisms. Standardized message formats and the implementation of context-aware communication approaches can greatly enhance both the clarity and relevance of exchanged information. Furthermore, integrating reinforcement or deep learning algorithms enables agents to adapt dynamically and develop cooperative behaviors over time. Applying these strategies in MAS results in more efficient and dependable performance within complex and ever-changing environments.

In a dynamic MAS environment, it is recommended to implement adaptive authentication mechanisms which adjust parameters according to the characteristics of messages and the context of interactions. For example, if a message appears to be suspicious, the system could require extra layers of authentication or verification. The research outlined in [18] introduces a solution which incorporates risk assessment with a platform-based MFA system. This approach can be seamlessly integrated into applications, enabling the delegation of the authentication process to an external resource while preserving the internal security of the system.

Combining various methods enhances the accuracy of the flow identification. For example, statistical analysis can be employed for preliminary data filtering [19], followed by the application of machine learning algorithms to improve classification accuracy [20].

MATERIALS AND METHODS

In order to address the challenge of identifying a specific message flow between two participants from a shared communication channel, a method which relies on analyzing the contextual characteristics of the shared message flow within the communication channel is proposed.

Using the previously referenced sensor network as an example of an MAS, the following characteristics can be identified as contextual for distinguishing message flows from multiple sensors at the gateway [21]:

- the frequency of sending messages serves as a quantitative parameter which measures the intensity of data transmission from individual sensor agents. It is determined by calculating the ratio between the number of messages sent by a sensor and a specified time interval.
- the message size represents the amount of data transmitted by the sensor, encompassing both the average and maximum sizes of messages sent. Given that sensors of the same type often transmit messages with uniform sizes, this characteristic can be used to identify and classify devices within the network.
- the message types categorize messages based on their content or purpose, such as requests, responses, or notifications. Since sensors of similar types typically send messages for analogous purposes, this parameter enables specific data flows from individual sources to be isolated and their activity evaluated.
- the send time refers to the time at which messages are transmitted, utilizing timestamps for analysis. This parameter makes it possible to detect activity patterns of sensors. However, data interpretation may require caution due to potential distortions caused by packet losses stemming from interference.
- the historical data comprises previous interaction records related to the source, including successful and failed authentication attempts.

The method is proposed, in order to address issues related to the collision of message identifiers. This approach relies on determining the source by counting the number of messages exchanged through the communication channel between two specific messages, thus testing the hypothesis that they originate from the same source. The method assumes that the contextual characteristics of agents remain relatively stable over time, enabling historical data to be used to analyze and distinguish message flows effectively. It takes into account that an originating agent may produce messages for a given target based on interactions with other agents, such as transmitting messages after a fixed or calculated number of exchanges with other MAS agents. This implies that the agent has the ability actively to manage the stability of its message flow contextual characteristics. The proposed solution offers notable benefits, including simplicity in implementation: a key consideration for devices with low computational capabilities and autonomous power supplies. Furthermore, this method can be executed independently on either the sender or the receiver side, since such meta-information does not need to be transmitted within service messages.

Several mechanisms can be employed to evaluate the statistical characteristics of message flow at the receiver during the processes of authentication and data flow identification:

1. Statistical analysis techniques such as clustering and regression analysis are effective in identifying patterns within the data and isolating the target flow.

Clustering allows for messages to be grouped around similar attributes, thus facilitating the distinction of message groups related to specific subscribers [22].

2. Machine learning algorithms, such as decision trees, neural networks, and support vector machines, can be trained on historical data to classify messages and identify the desired flow. For training purposes, structured data regarding interactions between specific subscribers can serve as a training dataset [23].
3. Time series analysis enables temporal dependencies in the data to be considered. This aids in the detection of anomalies which may signify the start or end of particular message flows. Techniques such as autoregressive models and moving average models can be applied to predict future messages based on past trends [24].

When relying on only one contextual characteristic, clustering is inadequate for decision-making due to the potential for a significant number of type I and type II errors, especially when similar characteristics appear in different flows. In addition, data derived from time series analysis can be heavily affected by the specific communication protocols in use. In fact, this has led the authors to abandon the continuous-time model [19], adopting instead a discrete-time approach based on counting messages transmitted in MAS. Consequently, a more suitable method for analyzing the resulting data involves machine learning. Logistic regression in this method serves as an efficient tool for reducing computational complexity during decision-making: an essential consideration for highly autonomous sensor devices in the MAS class under review.

A model is proposed for generating analyzed samples representing the number of messages transmitted in MAS between dataflow messages from a single source to a single receiver (Fig. 1).

This scenario contains a specific sequence of messages, $S_{i-1}, S_i, S_{i+1}, S_{i+l-1}, ..., S_{i+l}, S_{i+l+1}, S_{i+l+2}$, sent by the network target agent, alongside multiple messages, $\mathbf{D}_{i-1}, \mathbf{D}_i, \mathbf{D}_{i+1}, ..., \mathbf{D}_{i+k-1}, \mathbf{D}_{i+k}$, originating from other network agents and transmitted accordingly. The count of these messages from other agents within the system is measured between pairs S_{i-1} and S_i, S_i and S_{i+1}, ..., S_{i+l-1} and S_{i+l}. Let $n_{i-1}, n_i, n_{i+1}, ..., n_{i+k-1}, n_{i+k}$ denote the cardinalities of the associated sets. Using these values, a sequential dataset is constructed to represent the number of messages exchanged before, during, and after each message in the sequence sent by the target agent.

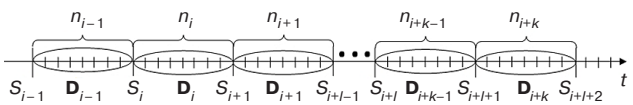

Fig. 1. A scheme for sampling statistical characteristics of the message flow depending on the message arrival time

Using fixed or calculated message values in MAS poses significant risks from an information security perspective. The inherent predictability allows an external observer to easily identify such message flows. As a consequence, the MAS structure becomes vulnerable and can be mapped by an attacker, even without directly analyzing the content of the transmitted data. However, introducing randomness in defining the sizes of sets $\mathbf{D}_{i-1}, \mathbf{D}_i, \mathbf{D}_{i+1}, ..., \mathbf{D}_{i+k-1}, \mathbf{D}_{i+k}$—where the distribution law and its parameters act as a shared secret between the sender and receiver—significantly enhances security. This approach complicates an attacker's ability to isolate messages from the source in the overall MAS message flow. Such increased unpredictability requires prolonged monitoring efforts, making unauthorized analysis much more challenging.

The message flow process operates as follows: once the source transmits a message, it waits for other MAS agents to send a random number of messages. This number can either remain constant or be calculated, as previously noted. In the study, a source is implemented where its messages are produced based on a count subordinate to the Poisson distribution. Randomizing the intervals allows for the messages from each specific source to be masked. This effectively conceals the MAS structure and the contextual details of the information flows from individual sources.

In the given scenario, unique message identifiers serve as verification sequences transmitted by a specific agent during a collision event (Fig. 2). Let us consider a sequence of messages, S_{i-1}, S_i, S_{i+1}, S_{i+l-1}, ..., S_{i+l}, S_{i+l+1}, S_{i+l+2}, generated by the target agent with inter-message intervals, $n_{i-1}, n_i, n_{i+1}, ..., n_{i+k-1}, n_{i+k}$, which are random values distributed according to the Poisson distribution in this scenario. Now, let us assume that the message S_{ab} is received between messages S_{i-1} and S_{i+1}. Upon verification of its identifier, it is determined to be a target agent-generated message appearing between two authentic messages. Since the unique parameters used to generate testing sequences are not known to the attacker, and any coincidence of these sequences in S_{ab} and S_i is entirely random, the message ID within the message sequence $n_{i-1} + n_i$ in MAS adheres to a random distribution. Therefore, the value of n_{ab} is uniformly distributed in the interval $n_{i-1} + n_i$.

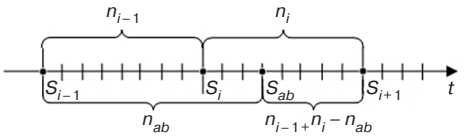

Fig. 2. Model for generating intervals in the case of a collision of unique verification sequences for two different messages

The task of the receiver involves analyzing two sets of paired values, (n_{i-1}, n_i) and $(n_{ab}, n_{i-1} + n_i - n_{ab})$, in order to calculate the probability that pair 1 follows the Poisson distribution, accounting for the generation of corresponding messages by the target agent. Meanwhile, pair 2 does not conform to this distribution, allowing message S_i to be excluded from consideration. These pairs can be enhanced by integrating a series of random values distributed according to the Poisson law. They represent the number of messages in MAS exchanged between different target source messages preceding or succeeding a collision. This approach considers not only a sequence of three messages where a collision takes place but also a broader sample of messages generated by the agent. However, the size of such a sample may not be large *a priori*. Expanding it excessively would lead to delays in determining which message caused the collision or require storing a substantial history of messages. Such scenarios would negatively impact the overall performance of the agents, especially their autonomy.

As highlighted in [25], skewness and kurtosis are effective tools for identifying the type of distribution. These coefficients, indicating variations in agent activity patterns, can serve as features for classifying message flows. However, when working with small sample sizes, the application of these coefficients requires caution. Their high sensitivity to extreme values makes them less reliable for small samples ($n < 50$), since standard coefficient errors tend to increase with decreasing n. This, in turn, diminishes the reliability of normality tests [26].

EXPERIMENTAL

The paper examines message flows in MAS to be identified using specific parameters for generating messages by the source and analyzing them at the receiver:

- M_chain = 11 refers to the number of messages from target agents analyzed, during which an ID collision occurred for one specific message.
- K = 70–110 represents the Poisson allocation parameter (accessible to the receiver), used by the source to define intervals between successive messages. In the context of MAS, this parameter corresponds to the ratio of message generation rates of all MAS agents to that of the source over a unit of time. It can also be understood as the total number of agents in MAS, assuming equal message generation rates among all agents.

Binary classification is used to assess the quality of the classification process using a message flow identification technique based on the individual statistical characteristics of each subscriber.

For each simulation cycle on the test dataset, the following steps are performed:

1. Create empty arrays to store information about messages and their properties.
2. Generate data:
   a) generate time intervals for the message sequence using the Poisson distribution (sequence A_1: $n_1, ..., n_9, n_{10}$).
   b) generate time intervals for the message sequence (sequence A_2: $n_1, ..., n_{ab}, n_9 + n_{10} - n_{ab}$, where n_{ab} is the evenly distributed number in the interval $n_9 + n_{10}$).
   c) calculate the skewness a and kurtosis e for two data sets, A_1 and A_2.
   d) assign labels to the input datasets, $f(e_1, a_1, e_2, a_2) = 1$, $f(e_2, a_2, e_1, a_1) = 0$, using labels from the interval $\{0, 1\}$.

Next, training and testing datasets are created from the generated dataset, in order to train the logistic regression model. Then the accuracy of the model classification on the testing dataset is assessed. Based on the values of type I and type II errors obtained, a classification matrix is constructed and ROC[3] analysis is performed to evaluate the model performance.

Logistic regression is used for binary classification. The likelihood of event Y occurring given the values of X is described by the following equation:

$$P(Y = 1 \mid X) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_n X_n)}}, \quad (1)$$

wherein $\beta_0, \beta_1, ..., \beta_n$ represent the model coefficients.

The following metrics are commonly employed to evaluate the performance of binary classification models:

1. Specificity (True Negative Rate) is calculated using the following formula:

$$TNR = \frac{TN}{TN + FP}, \quad (2)$$

wherein TN represents the number of true negatives while FP denotes the number of false positives.

2. Accuracy measures the degree to which the predicted outcomes match the actual results, expressed as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \quad (3)$$

wherein TP stands for true positives, FN for false negatives, TN for true negatives, and FP for false positives.

[3] Receiver operating characteristic.

3. Precision is described using the following formula:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \quad (4)$$

4. Recall (or sensitivity) is calculated as follows:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \quad (5)$$

5. $F1$ score is the harmonic mean of Precision and Recall, which is defined as follows:

$$F1 = 2 \cdot \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}. \quad (6)$$

The $F1$ score ranges from 0 to 1, where a score of 1 signifies perfect precision and recall, while a score of 0 indicates no correct predictions.

In multi-class classification tasks, the following approaches are commonly used to aggregate metrics like Precision, Recall, and $F1$ score:

1. Macro average is an average value associated with aggregated or general indicators. It is used to compare the performance of a model across all classes, regardless of the size of each class:

$$\text{Macro avg} = \frac{1}{N}\sum_{i=1}^{N} M_i, \quad (7)$$

wherein N represents the total number of classes while M_i is the metric value for class i.

2. Weighted average calculates the average metric value for each class, taking into account the number of instances in each class:

$$\text{Weight avg} = \frac{\sum_{i=1}^{N} M_i \omega_i}{\sum_{i=1}^{N} \omega_i}, \quad (8)$$

wherein ω_i represents the number of instances (or weight) for class i.

RESULTS AND DISCUSSION

The studies reveal that the metric values (Precision, Recall, $F1$) for each of the two classes (0; 1) with $K = 70$–110, vary between 0.81 and 0.85. These metrics are derived by analyzing the contextual characteristics of information flows from individual sources. A method is used which identifies message flows based on these contextual attributes. The Precision metric signifies that 81% to 85% of messages classified as belonging to the target source actually originate from it. Similarly, the Recall metric shows that 81% to 85% of all messages from the target source are correctly identified. The $F1$ score reflects the balance of the model, and demonstrates its effectiveness in minimizing both false positives and false negatives.

The Accuracy metric represents the overall percentage of messages correctly classified, indicating that 81% to 85% of all messages in the system (both target and irrelevant) are assigned to their correct categories. The alignment of the Macro average and Weighted average values suggests that there is no class imbalance. This is crucial because, in real-world scenarios, a balanced distribution of target and irrelevant messages makes analysis more straightforward and enhances method reliability.

The identical values for classes 0 and 1 can be attributed to the balanced data and the use of optimized feature selection, derived from the statistical characteristics of message flow.

Figure 3 illustrates the ROC curve for the logistic regression model at $K = 90$. The x-axis depicts the false positive rate, while the y-axis shows the true positive rate. The area under the curve (AUC) is calculated to be 0.83, suggesting that the model exhibits high sensitivity and effectively distinguishes between target and non-target messages, delivering satisfactory performance.

The experimental results indicate that the model operates reliably, even with a high value of K. Here, K represents the ratio of the total number of messages in MAS per unit of time to the number of messages from the target agent during the same time period.

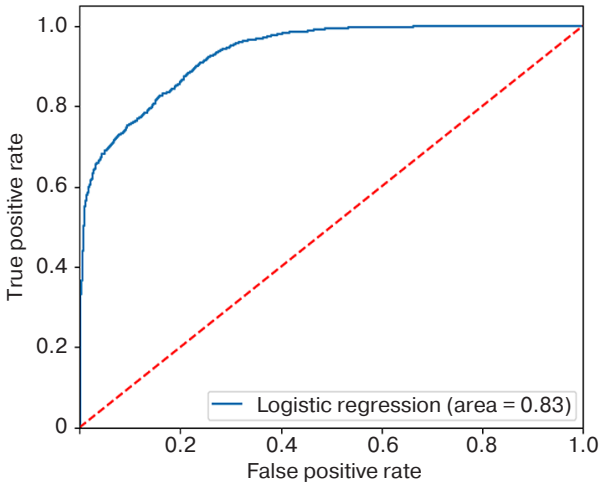

Fig. 3. ROC curve for the logistic regression model with $K = 90$ (solid blue line). The red dashed line represents a random classifier (AUC = 0.5)

Study [19], previously examined, introduces a methodology for identifying the origin of messages through statistical analysis of inter-packet interval times. It incorporates an evaluation of skewness and kurtosis, in which the defining rules are expressed as linear relationships between these parameters.

However, the study does not account for the impact of interference on the communication channel or delivery time. This omission is justified by disregarding the inter-packet interval time in the given context, allowing the influence of such interference to be reasonably overlooked.

The application of advanced statistical analysis, as opposed to the algebraic comparison of skewness and kurtosis, enables higher Precision metric values to be achieved than those reported in [19]. For example, with the same predefined message sequence length of M_chain = 11, a Precision metric in the range of 0.3–0.4 is achieved in [19].

When comparing the results of this study to the aforementioned method, it can be concluded that under comparable communication system parameters the Precision metric shows an improvement, increasing by 0.4–0.5.

CONCLUSIONS

The study of the hypothesis presented in this paper regarding the use of skewness and kurtosis in identifying the distribution type of random variables has demonstrated its potential for distinguishing differences in the activity patterns of agents based on their contextual characteristics.

Through simulation of the operation of the message flow identification subsystem in the event of a collision of unique message identifiers, satisfactory values for binary classification quality metrics have been achieved using a method for identifying message flows based on information flow characteristics from specific sources.

The approach described, which involves calculating the total number of messages exchanged between all MAS elements for a given target agent, is notable for its simplicity of implementation. This aspect is particularly important for low-performance devices which rely on autonomous power supplies. In addition, the method enables such counting to be performed independently on both the sender and receiver sides. The classification accuracy achieved by this approach ranges between 0.81 and 0.85.

The research findings demonstrate that leveraging contextual characteristics and statistical analysis facilitates accurate identification of target flows in MAS containing 70 to 110 agents. This method is especially suitable for low-bandwidth communication channels, in which it is crucial to minimize the size of transmitted batch headers and reduce the computational costs associated with authentication procedures. A particularly promising avenue for further development involves integrating adaptive mechanisms capable of handling dynamically changing traffic congestion, where the effective number of agents in MAS may vary widely.

Authors' contributions

M.O. Tanygin—concept development, methodology development, approval of the final version.

I.O. Mishin—conducting research, writing the draft, preparing and editing the text.

E.A. Kuleshova—conducting research, analyzing data, preparing and editing the text.

A.V. Kiselev—conducting research, analyzing data.

REFERENCES

1. Öztürk G., Saran N., Doğanaksoy A. Modified Attribute-Based Authentication for Multi-Agent Systems. *Int. J. Inform. Security Sci.* 2023;12(3):1–13. https://doi.org/10.55859/ijiss.1294580
2. Kuleshova E.A., Maruhlenko A.L., Dobritsa V.P., Tanygin M.O., Plugatorev A.V. A variant of the algorithm for generating pseudorandom binary sequences based on the properties of linear cellular automata. *Prikaspiiskii zhurnal: upravlenie i vysokie tekhnologii = Caspian Journal: Control and High Technologies.* 2021;54(2):62–70 (in Russ.). https://doi.org/10.21672/2074-1707.2021.53.1.062-070
3. Tanygin M.O., Kuleshova E.A., Mitrofanov A.V., Gladilina E.U. Increasing the speed of error detection when forming data block chains based on the analysis of the number of hash matches. *Prikaspiiskii zhurnal: upravlenie i vysokie tekhnologii = Caspian Journal: Control and High Technologies.* 2022;1(57):85–93 (in Russ.). https://doi.org/10.54398/2074-1707_2022_1_85
4. Yuan J., Yang J., Zhou S., Wang C. Efficient Group Authentication with Multiple Authentications on Resource-Limited Devices. *J. Supercomput.* 2025;81:929. https://doi.org/10.1007/s11227-025-07404-6
5. Gopirajan P.V., Mani K. Secure Multi-Authentication using BlockchainTechnology in Cloud based Internet of Things. *Telematique*. 2022;21(1):6640–6650.
6. Wee A.K., Chekole E.G., Zhou J. Excavating Vulnerabilities Lurking in Multi-Factor Authentication Protocols: A Systematic Security Analysis. *arXiv Cornell University*. 2024;2407(20459):1–24. https://doi.org/10.48550/arXiv.2407.20459
7. Chenchev I. Framework for Multi-factor Authentication with Dynamically Generated Passwords. In: Arai K. (Ed.). *Advances in Information and Communication. FICC 2023. Lecture Notes in Networks and Systems*. Springer; 2023. V. 652. P. 563–576. https://doi.org/10.1007/978-3-031-28073-3_39
8. Li H., Han D. Blockchain-assisted secure message authentication with reputation management for VANETs. *J. Supercomput.* 2023;79(17):19903–19933. https://doi.org/10.1007/s11227-023-05394-x
9. Xu Y., Jian X., Li T., Zou S., Li B. Blockchain-Based Authentication Scheme with an Adaptive Multi-Factor Authentication Strategy. *Mobile Inform. Syst.* 2023;2023:4764135. https://doi.org/10.1155/2023/4764135

10. Liu J., Mu Q., Che R., et al. Multi-participant quantum anonymous communication based on high-dimensional entangled states. *Physica Scripta*. 2024;99(9):095109. https://doi.org/10.1088/1402-4896/ad69d9
11. Kuleshova E.A., Tanygin M.O. Study of characteristics of modern generators of pseudorandom sequences. *Telekommunikatsii = Telecommunications*. 2023;7:28–39 (in Russ.). https://doi.org/10.31044/1684-2588-2023-0-7-28-39
12. Tanygin M.O. Restoring the order of information packets based on hash sequence analysis. *Izvestiya Yugo-Zapadnogo gosudarstvennogo universiteta = Proceedings of the Southwest State University*. 2020;24(1):175–188 (in Russ.).
13. Alamgir N., Negati S., Bright C. SHA-256 Collision Attack with Programmatic SAT. *arXiv Cornell University.* 2024;2406.20072. https://doi.org/10.48550/arXiv.2406.20072
14. Fang Y. A research on different digital signature schemes. *Appl. Comput. Eng.* 2023;16(1):27–35. http://doi.org/10.54254/2755-2721/16/20230855
15. Phatangare S., Jadhav S., Kawane S., Holkar P., Gaikwad P. Multi-Level Encryption System using AES and RSA Algorithms. *Int. J. Res. Appl. Sci. Eng. Technol.* 2024;15(5):4043–4051. https://doi.org/10.22214/IJRASET.2024.62420
16. Tanygin M.O., Alshaeaa H.Y., Kuleshova E.A. A method of the transmitted blocks information integrity control. *Radio Electronics, Computer Science, Control.* 2020;1:181–189.
17. Tao M., Li Q., Yu J. Multi-Objective Dynamic Path Planning with Multi-Agent Deep Reinforcement Learning. *J. Marin. Sci. Eng.* 2025;13(1):20. https://doi.org/10.3390/jmse13010020
18. Morais D., Zuquete A., Mendes A. Adaptive, Multi-Factor Authentication as a Service for Web Applications. In: *2023 7th Cyber Security in Networking Conference* (*CSNet*). 2023. P. 74–80. http://doi.org/10.1109/CSNet59123.2023.10339695
19. Plugatarev A.V. Model for determining the message source by statistical analysis of metadata in an open communication channel. *Prikaspiiskii zhurnal: upravlenie i vysokie tekhnologii = Caspian Journal: Control and High Technologies.* 2022;4(60):30–37 (in Russ.).
20. Dharrao D., Gaikwad P., Gawai S.V., Bongale A.M., Patel K., Singh A. Classifying SMS as spam or ham: Leveraging NLP and machine learning techniques. *Int. J. Saf. Secur. Eng.* 2024;14(1):289–296. https://doi.org/10.18280/ijsse.140128
21. Placzek B. A Multi-Agent Prediction Method for Data Sampling and Transmission Reduction in Internet of Things Sensor Networks. *Sensors*. 2023;23(20):8478. https://doi.org/10.3390/s23208478
22. Vedmiediev D., Shapoval N. Text Message Clustering. *Electronics and Control Systems.* 2023;4(78):16–20.
23. Katwal S., Sharma N., Kumar K. A Deep Learning Approach for Throughput Enhanced Clustering and Spectrally Efficient Resource Allocation in Ultra-Dense Networks. *IEEE Trans. Netw. Service Manag*. 2025;22(1):582–591. https://doi.org/10.1109/TNSM.2024.3470235
24. Huang X., Zhou S. QMNet: Importance-Aware Message Exchange for Decentralized Multi-Agent Reinforcement Learning. *IEEE Trans. Mobile Comput*. 2023;23(5):4739–4751. https://doi.org/10.1109/TMC.2023.3296726
25. Goloveshkin V.A., Zhukova G.N., Ulyanov M.V., Fomichev M.I. The estimation of the complexity of solving a particular travelling salesman problem by quantile-based measures for skewness and kurtosis. *Int. J. Open Inform. Technol.* 2016;4(12):7–12 (in Russ.). https://elibrary.ru/xetabh
26. Tanygin M.O., Dobritsa V.P., Mitrofanov A.V., Ahmat Kh.I. Mathematical interpretation of the results of cognitive analysis of network packets metadata. *Izvestiya Yugo-Zapadnogo gosudarstvennogo universiteta = Proceedings of the Southwest State University*. 2023;27(3):66–78 (in Russ.).

СПИСОК ЛИТЕРАТУРЫ

1. Öztürk G., Saran N., Doğanaksoy A. Modified Attribute-Based Authentication for Multi-Agent Systems. *Int. J. Inform. Security Sci.* 2023;12(3):1–13. https://doi.org/10.55859/ijiss.1294580
2. Кулешова Е.А., Марухленко А.Л., Добрица В.П., Таныгин М.О., Плугаторев А.В. Вариант алгоритма генерации псевдослучайных двоичных последовательностей, основанный на свойствах линейных клеточных автоматов. *Прикаспийский журнал: управление и высокие технологии.* 2021;54(2):62–70. https://doi.org/10.21672/2074-1707.2021.53.1.062-070
3. Таныгин М.О., Кулешова Е.А., Митрофанов А.В., Гладилина Е.Ю. Повышение скорости обнаружения ошибок при формировании цепочек блоков данных на основе анализа числа совпадений хешей. *Прикаспийский журнал: управление и высокие технологии.* 2022;1(57):85–93. https://doi.org/10.54398/2074-1707_2022_1_85
4. Yuan J., Yang J., Zhou S., Wang C. Efficient Group Authentication with Multiple Authentications on Resource-Limited Devices. *J. Supercomput.* 2025;81:929. https://doi.org/10.1007/s11227-025-07404-6
5. Gopirajan P.V., Mani K. Secure Multi-Authentication using BlockchainTechnology in Cloud based Internet of Things. *Telematique*. 2022;21(1):6640–6650.
6. Wee A.K., Chekole E.G., Zhou J. Excavating Vulnerabilities Lurking in Multi-Factor Authentication Protocols: A Systematic Security Analysis. *arXiv Cornell University*. 2024;2407(20459):1–24. https://doi.org/10.48550/arXiv.2407.20459
7. Chenchev I. Framework for Multi-factor Authentication with Dynamically Generated Passwords. In: Arai K. (Ed.). *Advances in Information and Communication. FICC 2023. Lecture Notes in Networks and Systems*. Springer; 2023. V. 652. P. 563–576. https://doi.org/10.1007/978-3-031-28073-3_39
8. Li H., Han D. Blockchain-assisted secure message authentication with reputation management for VANETs. *J. Supercomput.* 2023;79(17):19903–19933. https://doi.org/10.1007/s11227-023-05394-x
9. Xu Y., Jian X., Li T., Zou S., Li B. Blockchain-Based Authentication Scheme with an Adaptive Multi-Factor Authentication Strategy. *Mobile Inform. Syst.* 2023;2023:4764135. https://doi.org/10.1155/2023/4764135
10. Liu J., Mu Q., Che R., et al. Multi-participant quantum anonymous communication based on high-dimensional entangled states. *Physica Scripta*. 2024;99(9):095109. https://doi.org/10.1088/1402-4896/ad69d9

11. Кулешова Е.А., Таныгин М.О. Исследование характеристик современных генераторов псевдослучайных последовательностей. *Телекоммуникации*. 2023;7:28–39. https://doi.org/10.31044/1684-2588-2023-0-7-28-39
12. Таныгин М.О. Восстановление порядка следования информационных пакетов на основе анализа хеш-последовательностей. *Известия Юго-Западного государственного университета*. 2020;24(1):175–188.
13. Alamgir N., Negati S., Bright C. SHA-256 Collision Attack with Programmatic SAT. *arXiv Cornell University.* 2024;2406.20072. https://doi.org/10.48550/arXiv.2406.20072
14. Fang Y. A research on different digital signature schemes. *Appl. Comput. Eng.* 2023;16(1):27–35. http://doi.org/10.54254/2755-2721/16/20230855
15. Phatangare S., Jadhav S., Kawane S., Holkar P., Gaikwad P. Multi-Level Encryption System using AES and RSA Algorithms. *Int. J. Res. Appl. Sci. Eng. Technol.* 2024;15(5):4043–4051. https://doi.org/10.22214/IJRASET.2024.62420
16. Tanygin M.O., Alshaeaa H.Y., Kuleshova E.A. A method of the transmitted blocks information integrity control. *Radio Electronics, Computer Science, Control.* 2020;1:181–189.
17. Tao M., Li Q., Yu J. Multi-Objective Dynamic Path Planning with Multi-Agent Deep Reinforcement Learning. *J. Marin. Sci. Eng.* 2025;13(1):20. https://doi.org/10.3390/jmse13010020
18. Morais D., Zuquete A., Mendes A. Adaptive, Multi-Factor Authentication as a Service for Web Applications. In: *2023 7th Cyber Security in Networking Conference* (*CSNet*). 2023. P. 74–80. http://doi.org/10.1109/CSNet59123.2023.10339695
19. Плугатарев А.В. Модель определения источника сообщений на основе статистического анализа метаданных в открытом канале связи. *Прикаспийский журнал: управление и высокие технологии.* 2022;4(60):30–37.
20. Dharrao D., Gaikwad P., Gawai S.V., Bongale A.M., Patel K., Singh A. Classifying SMS as spam or ham: Leveraging NLP and machine learning techniques. *Int. J. Saf. Secur. Eng.* 2024;14(1):289–296. https://doi.org/10.18280/ijsse.140128
21. Placzek B. A Multi-Agent Prediction Method for Data Sampling and Transmission Reduction in Internet of Things Sensor Networks. *Sensors*. 2023;23(20):8478. https://doi.org/10.3390/s23208478
22. Vedmiediev D., Shapoval N. Text Message Clustering. *Electronics and Control Systems.* 2023;4(78):16–20.
23. Katwal S., Sharma N., Kumar K. A Deep Learning Approach for Throughput Enhanced Clustering and Spectrally Efficient Resource Allocation in Ultra-Dense Networks. *IEEE Trans. Netw. Service Manag*. 2025;22(1):582–591. https://doi.org/10.1109/TNSM.2024.3470235
24. Huang X., Zhou S. QMNet: Importance-Aware Message Exchange for Decentralized Multi-Agent Reinforcement Learning. *IEEE Trans. Mobile Comput*. 2023;23(5):4739–4751. https://doi.org/10.1109/TMC.2023.3296726
25. Головешкин В.А., Жукова Г.Н., Ульянов М.В., Фомичев М.И. Использование квантильных коэффициентов асимметрии и эксцесса для оценки сложности решения задачи коммивояжера. *Int. J. Open Inform. Technol.* 2016;4(12):7–12. https://elibrary.ru/xetabh
26. Таныгин М.О., Добрица В.П., Митрофанов А.В., Ахмат Х.И. Математическая интерпретация результатов когнитивного анализа метаданных сетевых пакетов. *Известия Юго-Западного государственного университета.* 2023;27(3):66–78.

About the Authors

Maxim O. Tanygin, Dr. Sci. (Eng.), Associate Professor, Dean of the Faculty of Fundamental and Applied Informatics, Southwest State University (94, 50 Let Oktyabrya ul., Kursk, 305040 Russia). E-mail: tanygin@yandex.ru. Scopus Author ID 19640649200, ResearcherID N-7689-2016, RSCI SPIN-code 2639-4800, https://orcid.org/0000-0002-4099-1414

Ilya O. Mishin, Postgraduate Student, Department of Information Security, Southwest State University (94, 50 Let Oktyabrya ul., Kursk, 305040 Russia). E-mail: mishin.ilya46@yandex.ru. ResearcherID MXJ-7912-2025, RSCI SPIN-code 6911-3642, https://orcid.org/0009-0006-8883-1731

Elena A. Kuleshova, Cand. Sci. (Eng.), Associate Professor, Department of Information Security, Southwest State University (94, 50 Let Oktyabrya ul., Kursk, 305040 Russia). E-mail: lena.kuleshova.94@mail.ru. Scopus Author ID 57216349335, ResearcherID AAI-9214-2021, RSCI SPIN-code 9607-8582, https://orcid.org/0000-0002-8270-564X

Alexey V. Kiselev, Cand. Sci. (Eng.), Associate Professor, Computer Engineering Department, Southwest State University (94, 50 Let Oktyabrya ul., Kursk, 305040 Russia). E-mail: kiselevalexey1990@gmail.com. Scopus Author ID 57337411000, ResearcherID S-9914-2018, RSCI SPIN-code 2016-7550, https://orcid.org/0000-0001-7228-0281

Об авторах

Таныгин Максим Олегович, д.т.н., доцент, декан факультета фундаментальной и прикладной информатики, ФГБОУ ВО «Юго-Западный государственный университет» (305040, Россия, Курск, ул. 50 лет Октября, д. 94). E-mail: tanygin@yandex.ru. Scopus Author ID 19640649200, ResearcherID N-7689-2016, SPIN-код РИНЦ 2639-4800, https://orcid.org/0000-0002-4099-1414

Мишин Илья Олегович, аспирант, кафедра информационной безопасности, ФГБОУ ВО «Юго-Западный государственный университет» (305040, Россия, Курск, ул. 50 лет Октября, д. 94). E-mail: mishin.ilya46@yandex.ru. ResearcherID MXJ-7912-2025, SPIN-код РИНЦ 6911-3642, https://orcid.org/0009-0006-8883-1731

Кулешова Елена Александровна, к.т.н., доцент, кафедра информационной безопасности, ФГБОУ ВО «Юго-Западный государственный университет» (305040, Россия, Курск, ул. 50 лет Октября, д. 94). E-mail: lena.kuleshova.94@mail.ru. Scopus Author ID 57216349335, ResearcherID AAI-9214-2021, SPIN-код РИНЦ 9607-8582, https://orcid.org/0000-0002-8270-564X

Киселев Алексей Викторович, к.т.н., доцент, кафедра вычислительной техники, ФГБОУ ВО «Юго-Западный государственный университет» (305040, Россия, Курск, ул. 50 лет Октября, д. 94). E-mail: kiselevalexey1990@gmail.com. Scopus Author ID 57337411000, ResearcherID S-9914-2018, SPIN-код РИНЦ 2016-7550, https://orcid.org/0000-0001-7228-0281

Translated from Russian into English by Kirill V. Nazarov
Edited for English language and spelling by Dr. David Mossop