UDC 004.056

DATA PROTECTION IN CLOUD SYSTEMS USING MACHINE LEARNING-BASED ENCRYPTION

Davydova A.
bachelor's degree, Siberian Federal University (Krasnoyarsk, Russia)

ЗАЩИТА ДАННЫХ В ОБЛАЧНЫХ СИСТЕМАХ С ИСПОЛЬЗОВАНИЕМ ШИФРОВАНИЯ НА ОСНОВЕ МАШИННОГО ОБУЧЕНИЯ

Давыдова А.С.
бакалавр, Сибирский федеральный университет (Красноярск, Россия)

Abstract

This article explores modern data encryption methods in cloud systems using machine learning (ML) algorithms. The main approaches include adaptive encryption models that adjust data protection processes based on the data type and volume. Special attention is given to ML-based anomaly detection methods, enabling prompt identification of potential threats and preventing data leaks. The use of hybrid encryption methods, combining symmetric and asymmetric encryption, ensures high security levels and optimized computational costs. Emphasis is placed on ML's role in enhancing encryption key distribution and risk prediction processes. Practical recommendations for implementing proposed solutions aimed at improving cloud system reliability are presented. The analysis shows that the use of ML in cloud data encryption significantly enhances data security, minimizes risks, and boosts cloud technology performance.

Keywords: data encryption, cloud systems, machine learning, anomaly detection, hybrid encryption.

Аннотация

В статье рассматриваются современные методы шифрования данных в облачных системах с использованием алгоритмов машинного обучения (МО). Приведены основные подходы, включающие адаптивные модели шифрования, позволяющие настраивать процесс защиты данных в зависимости от их типа и объема. Особое внимание уделено методам обнаружения аномалий на основе МО, которые позволяют оперативно выявлять потенциальные угрозы и предотвращать утечки информации. Рассматривается применение гибридных методов шифрования, сочетающих симметричное и асимметричное шифрование, обеспечивающих высокий уровень безопасности и оптимизацию вычислительных затрат. Подчеркивается значимость МО для улучшения процессов распределения ключей шифрования и прогнозирования рисков. Представлены рекомендации по практическому применению предложенных решений, направленных на повышение надежности облачных систем. Проведенный анализ демонстрирует, что использование МО в шифровании данных в облаке значительно улучшает защиту данных, минимизирует риски и повышает производительность облачных технологий.

Ключевые слова: шифрование данных, облачные системы, машинное обучение, обнаружение аномалий, гибридное шифрование.

Introduction

Modern cloud systems, which provide data storage and processing, offer companies and users convenient access to information, significantly expanding the possibilities for interaction with digital resources. However, as the use of cloud technologies grows, so do the risks associated with data security, especially the threats of unauthorized access, theft, and data leaks. In this context, one of the most in-demand methods for protecting information in cloud systems is data encryption, which secures information from potential attacks and guarantees its confidentiality. At the same time, traditional encryption methods often encounter difficulties when working with large volumes of data typical for cloud systems. This leads to the need for developing new approaches that can not only ensure protection but also optimize the encryption and decryption processes. Machine learning (ML) opens up opportunities for creating adaptive encryption methods that can automatically adjust to the type and volume of data, as well as the parameters affecting security and performance. Such methods allow for minimizing computational costs and improving security metrics [1].

The purpose of this article is to explore approaches to data encryption in cloud systems using ML algorithms. The article will analyze modern ML-based methods for encrypting and decrypting data, as well as provide examples and recommendations for their application to enhance security levels in cloud environments.

Main part. Application of ML for optimizing encryption

Traditional encryption methods, while reliable, often face limitations when working with large volumes of data, which is characteristic of cloud systems. To overcome these challenges, ML offers adaptive encryption models that can adjust the encryption process based on the specifics and volume of data [2, 3]. Machine learning algorithms, such as neural networks, can be trained on analyzing data streams and selecting the most effective parameters for encryption, thereby reducing computational resource costs and increasing performance.

Figure 1 illustrates the relationship between the volume of data and the time required for encryption. This demonstrates how larger datasets typically lead to longer encryption times, highlighting the need for optimization.

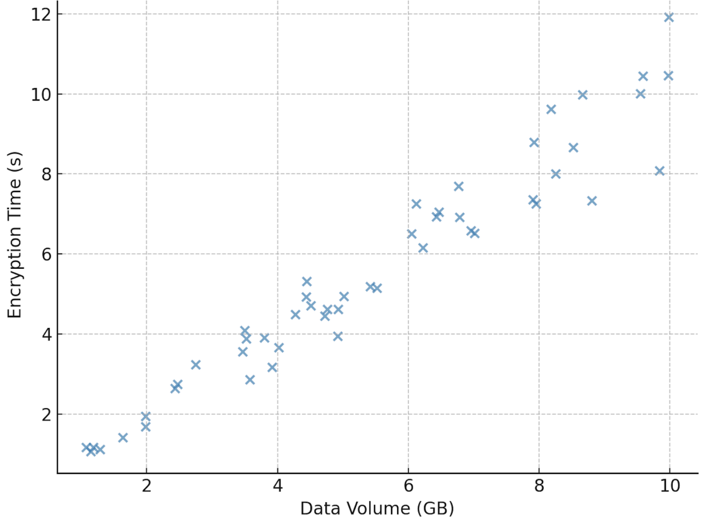

Figure 1. Dependency between data volume and encryption time

As shown in Figure 1, there is a clear trend indicating that as the data volume increases, the encryption time rises proportionally. This emphasizes the importance of adaptive ML-based encryption methods, which can minimize computational costs and enhance performance. By analyzing data streams and dynamically adjusting encryption parameters, these methods can significantly improve efficiency, even for large-scale cloud systems. ML can also be used to improve the distribution of encryption keys in cloud systems. ML methods allow for forecasting temporal patterns of data usage, which aids in timely key updates and reduces the risk of attacks. For example, clustering methods help categorize data streams, highlighting data that requires a higher level of protection. This optimizes resource usage by distributing the load on encryption systems.

Protection from attacks through anomaly analysis

ML is also applied to detect anomalies in data streams, enabling prompt identification of potential threats. For instance, anomaly detection algorithms can analyze network traffic and user behavior, identifying unusual actions that may signal attempts at hacking or unauthorized access.

Algorithms such as isolation forests and autoencoders analyze habitual activity patterns, and when deviations occur, checks and additional security measures are triggered [4]. Using ML in these processes allows the system to self-learn from incoming data and improve the accuracy of threat detection, adapting to changing conditions. This approach enhances the reliability of the system and minimizes the risks of data loss by applying preventive measures based on analysis results.

Application of hybrid encryption methods for cloud data

Hybrid encryption methods, which combine the advantages of symmetric and asymmetric approaches, are also successfully implemented in cloud systems. For instance, during data transmission, asymmetric methods may be used for secure key transfer, while symmetric algorithms encrypt the content, which reduces processing time and maintains a high level of security [5]. Combined with ML, hybrid methods can adapt to the type of data and transmission conditions.

The hybrid approach using ML allows for selecting the most appropriate encryption method for specific tasks, such as protecting databases, file storage, or messaging systems. ML analyzes the volume, nature, and frequency of data transmission and can predict the best parameters for encryption, thereby maintaining a balance between security and performance.

Application of encryption in cloud systems

Data encryption in cloud systems is one of the key methods of ensuring security, as it protects data from unauthorized access at all stages of its storage and transmission [6]. Data stored in the cloud can be subject to various attacks, including interception during transmission and attacks on storage servers. Encrypting data at both the server and client levels helps prevent information leaks, as even if attackers gain access to encrypted data without the decryption key, it remains unreadable.

Several stages play an important role in encryption. First, encrypting data before it is sent to the cloud, known as client-side encryption, ensures that the information is protected before it leaves the user's device. Thus, only encrypted data reaches the cloud, and neither the cloud provider nor third parties can access the original data [7]. Second, server-side encryption is applied when storing data in the cloud, protecting it from unauthorized access within the cloud infrastructure itself. The third stage is encrypting data during transmission, which protects data from interception when exchanged between the client and server or between servers.

Cloud systems widely employ encryption methods such as symmetric and asymmetric encryption, each having its own advantages and disadvantages [8]. For example, symmetric encryption uses a single key for both encrypting and decrypting data, making it faster but less secure in terms of key storage. Asymmetric encryption uses two different keys – a public key and a private key – enhancing security but requiring greater computational costs.

Conclusion

Thus, data encryption in cloud systems using ML methods represents an innovative approach to ensuring information security, which is particularly relevant given the growing volume of data and heightened risks of cyberattacks. ML not only optimizes encryption processes but also adapts them to dynamic conditions, reducing computational costs and improving system performance. The application of adaptive encryption models and anomaly detection technologies based on ML enables timely identification of suspicious activity, significantly reducing the likelihood of unauthorized access and data leaks. ML methods can be used to create hybrid solutions that combine the best features of symmetric and asymmetric encryption algorithms, ensuring a high level of confidentiality and reliability.

Therefore, the combination of traditional encryption methods and modern ML approaches allows for the creation of a multi-layered data protection system in the cloud that can withstand modern threats. Further research and implementation of these methods will help ensure data security and increase user trust in cloud technologies, opening new prospects for their use in various industries.

References

1. Krasov A.V., Shterenberg S.I., Fakhrutdinov R.M., Ryzhakov D.V., Pestov I.E. Analysis of information security of an enterprise based on collecting user data from open sources and monitoring

information resources using machine learning // T-Comm – Telecommunications and Transport. 2018. Vol. 12. No.10. P. 36-40.
2. Kalinine M.O., Shterenberg S.I. Analysis of information security of an enterprise based on monitoring information resources using machine learning // Intelligent technologies in transport. 2018. No.3(15). P. 47-54.
3. Vavilova A.S. Overview of existing mechanisms for ensuring data confidentiality in machine learning-based systems // Innovations. Science. Education. 2021. No.36. P. 1159-1167.
4. Babenko L.K., Shumilin A.S., Alexeyev D.M. Algorithm for ensuring the protection of confidential data in a cloud medical information system // News of the Southern Federal University. Technical sciences. 2021. No.5(222). P. 120-134.
5. Kuznetsov I.A. Security and confidentiality of data in mobile applications developed using machine learning technologies // Cold Science. 2024. No.2. P. 5-13.
6. Bespalova N.V., Nechaev S.V. Ensuring information security in cloud storage // Security issues. 2023. No.2. P. 19-26.
7. Vlasov R.S., Lukashik E.P., Osipyan V.O. Development of a mechanism for secure communication based on machine learning methods // Caspian Journal: Management and High Technologies. 2020. No.2(50). P. 108-117.
8. Angapov V.D. Overview of modern cloud platforms for machine learning purposes // Problems of modern science and education. 2023. No.7(185). P. 5-17.