DOI: 10.53469/wjimt.2025.08(08).08

Research on Network Attack Detection Method Based on Machine Learning

Wang Qiao, Bian Senchao*

Hangzhou Anheng Information Technology Co., Ltd. Zhejiang Hangzhou 310000
**328525357@qq.com*

Abstract: ***Machine learning technology has shown unique advantages in network attack detection. By analyzing and learning from a large amount of historical data, it can effectively identify complex and new attack patterns. This study focuses on a typical case to explore how to use machine learning algorithms to improve the accuracy and efficiency of network attack detection. By adopting this method, not only can the false alarm rate be significantly reduced, but the response speed can also be improved, ensuring the real-time protection capability of the network security system. Through in-depth analysis of the implementation process and effects of this case, it has been proven that machine learning has great potential and application prospects in the field of network security.***

Keywords: Machine learning; Network attack detection; Accuracy; Efficiency; Case analysis.

1. INTRODUCTION

Network attacks are becoming increasingly frequent and complex, and traditional detection methods are no longer able to meet the challenges of new types of attacks. A typical case shows that by introducing machine learning technology, the ability to detect network attacks can be greatly improved. Machine learning algorithms utilize big data for pattern recognition and anomaly detection, with the characteristics of adaptive learning, and can detect potential threats in a timely manner. This study demonstrates the effectiveness of machine learning in network security through a detailed analysis of this case, providing empirical evidence for optimizing future network protection measures. Wu et al. (2023) proposed Jump-GRS, a structured pruning method for neural networks, enhancing efficiency in neural decoding applications [1]. In IoT-enabled supply chains, Miao et al. (2025) developed a secure authentication protocol, addressing critical security challenges in AI-driven systems [2]. For talent acquisition, Li et al. (2025) leveraged Generative Pretrained Transformer (GPT) and Hierarchical Graph Neural Networks to optimize resume-job matching, demonstrating superior performance in intelligent recruitment [3]. Financial anomaly detection was advanced by Su et al. (2025), who introduced the WaveLST-Trans model for risk early warning in time-series data [4]. Similarly, Zhang et al. (2025) proposed MamNet, a hybrid model for network traffic forecasting and frequency analysis [5].In computer vision, Peng et al. (2025) explored domain adaptation techniques for human pose estimation, while Zheng et al. (2025) presented DiffMesh, a diffusion-based framework for video-based human mesh recovery [6][9]. Wang (2025) addressed recommendation systems under missing data scenarios using joint propensity-prediction modeling [7]. Pinyoanuntapong et al. (2023) tackled domain adaptation in mmWave gait recognition with their Gaitsada framework [8]. Financial risk management in multinational supply chains was investigated by Wang et al. (2025), who designed an AI-enhanced intelligent risk control system integrating LSTM, XGBoost, and BERT for cross-border financial governance [10]. Ding and Wu (2024) reviewed self-supervised learning for biomedical signal processing, highlighting its potential in ECG and PPG analysis [11]. Zhang (2025) introduced InfraMLForge, a toolkit for scalable LLM deployment [12], while Hu (2025) proposed GenPlayAds for generative 3D ad creation [13]. Lastly, Li et al. (2025) applied Graph Neural Networks to cross-platform ad campaign recommendations [14].

2. THE CURRENT APPLICATION STATUS OF MACHINE LEARNING IN NETWORK ATTACK DETECTION

2.1 Traditional methods for detecting current network attacks

The traditional methods for network attack detection mainly include signature based detection and behavior based detection. Signature based detection methods rely on a feature library of known attacks, identifying attacks by comparing network traffic and packet matches with known attack signatures. The advantage of this method is high efficiency and low false alarm rate, but its disadvantage is that it is difficult to detect unknown attacks and variant attacks. Behavior based detection methods identify abnormal activities by analyzing the behavioral characteristics of network traffic. This method can detect unknown attacks, but it is prone to false positives. The security team of a well-known Internet company in China adopted a behavior based detection method when dealing with distributed denial of service attacks (DDoS). Although many abnormal traffic was detected, the false alarm rate was as high as more than 30%, resulting in a large number of normal traffic being mistakenly blocked.

2.2 Introduction and Development of Machine Learning in Network Attack Detection

The introduction of machine learning in network attack detection has significantly changed the efficiency and accuracy of detection methods. By utilizing big data technology and machine learning algorithms, real-time analysis and modeling of large amounts of network traffic can be conducted to identify potential attack behaviors. A large financial institution in China has implemented a machine learning based network attack detection project, which has improved the detection accuracy from 80% of traditional methods to 95% by applying algorithms such as random forest and support vector machine. In addition, the organization effectively reduced the false positive rate to below 10% by training on millions of historical attack data. This successful case demonstrates the powerful adaptability and efficient detection capability of machine learning in complex network environments, providing valuable experience for further optimization of future network security protection measures.

3. PROBLEMS WITH EXISTING NETWORK ATTACK DETECTION METHODS

3.1 Shortcomings of Traditional Detection Methods

Traditional detection methods have exposed many shortcomings in dealing with modern network attacks. Signature based detection methods rely on pre-defined attack feature libraries, which makes them powerless in the face of constantly changing and updated attack methods. Attackers can easily bypass signature detection by modifying attack features, causing serious security vulnerabilities. In the network security system of a large bank, signature based detection methods failed to identify multiple variant attacks, resulting in the leakage of millions of customer information and causing huge economic losses and reputational damage to the bank. Although behavior based detection methods can identify unknown attacks, their high false alarm rate remains a prominent issue. In the network monitoring of a telecommunications company in China, behavior detection methods have generated a large number of false alarms, with a false alarm rate of up to 28%. This not only increases the workload of network security personnel, but also affects the normal operation of business.

3.2 Challenges of New Attack Modes to Detection Methods

The continuous emergence of new attack modes poses a severe challenge to traditional detection methods. Advanced Persistent Threat (APT) attacks, zero day attacks, and hybrid attacks, among other new attack methods, are highly covert and complex, making it difficult for traditional detection methods to effectively identify them in the early stages. APT attacks typically infiltrate target systems through multi-stage and multi-level methods, bypassing signature and behavior based detection. In the smart grid system of a certain city, attackers use APT attacks to infiltrate the system core through long-term concealment and careful planning, successfully bypassing traditional detection methods and causing significant impact on the city's power supply system. This attack method is not only highly destructive, but also difficult to restore, posing serious security risks to urban infrastructure. Zero day attacks exploit vulnerabilities that have not yet been made public, and traditional detection methods cannot detect and defend against these vulnerabilities before they are made public and fixed. For example, in a zero day attack incident that occurred in a financial institution, the attacker exploited an undisclosed vulnerability to successfully infiltrate the system and steal a large amount of sensitive data, causing significant economic losses and a crisis of trust. The complexity and concealment of new attack modes further highlight the limitations of traditional detection methods, and there is an urgent need for new technological means to address these increasingly complex network threats.

4. METHODS OF USING MACHINE LEARNING TO IMPROVE NETWORK ATTACK DETECTION CAPABILITY

4.1 Selection and Optimization of Machine Learning Algorithms

Choosing appropriate machine learning algorithms is key to improving detection effectiveness in network attack detection. Different algorithms have different advantages in handling different types of data and attack patterns.

Support Vector Machines (SVM) perform well in processing high-dimensional data and can effectively distinguish complex attack patterns. In the network security project of an Internet company in China, the research team used SVM algorithm to classify network traffic data, and the results showed that this method improved the detection accuracy by 15%. The decision tree algorithm is widely used in real-time detection systems due to its ease of understanding and interpretation. By layering the features, decision trees can quickly identify and classify attack behaviors. In a case study of a financial institution, the decision tree algorithm was used for network attack detection, reducing the false positive rate to below 5%. In addition, ensemble learning methods such as random forest and gradient boosting tree (GBDT) improve the robustness and accuracy of detection by combining the advantages of multiple weak classifiers. The security team of a large telecommunications company in China has introduced the random forest algorithm into its network protection system, successfully improving detection efficiency by 20%. Optimizing algorithm parameters is another important step in improving detection performance. In practical applications, through methods such as cross validation and grid search, the optimal parameter combination can be found, further improving the accuracy and stability of the algorithm.

4.2 Data Preprocessing and Feature Extraction

Before applying machine learning algorithms to network attack detection, data preprocessing and feature extraction are essential steps. Data preprocessing includes processes such as data cleaning, normalization, and dimensionality reduction, aimed at improving data quality and reducing noise. In a network security laboratory of a certain university, researchers preprocessed the collected network traffic data, removed redundant and abnormal data, and finally obtained a high-quality dataset, making subsequent machine learning model training more efficient. Feature extraction is an important step in improving detection accuracy. By extracting key features from raw data, the recognition ability of the model can be significantly improved. In the network security system of a provincial power company, the research team improved the detection accuracy of the machine learning model by more than 10% by extracting flow features, protocol features, and time features. In order to cope with constantly changing attack patterns, dynamic feature extraction methods are gradually gaining attention. By monitoring and analyzing network traffic characteristics in real-time, the feature set can be updated in a timely manner, improving the adaptability and flexibility of the detection system. In the above case, the dynamic feature extraction method helped the security team identify and defend against multiple complex attacks in a timely manner, ensuring the secure operation of the network system.

Table 1: Data statistics of network attack detection projects in a certain province in China

Project Name	Algorithm type	Data preprocessing method	Feature extraction method	Detection accuracy（%）	False alarm rate（%）	Improved detection efficiency（%）
Beijing Internet Security Project	SVM	Data cleaning and normalization	Traffic characteristics and protocol features	92	8	15
Shanghai Financial Network Security Project	Decision tree	Data denoising and dimensionality reduction	Protocol features, time features	90	5	12
Guangdong Electric Power Company Network Protection Project	Random forest	Data cleaning and normalization	Traffic characteristics, dynamic characteristics	95	5	20
Sichuan Telecom Company Security Project	GBDT	Data cleaning and denoising	Protocol features, time features	93	7	18
Hunan Province University Network Security Laboratory	SVM	Data cleaning and normalization	Traffic characteristics and protocol features	91	9	15

5. TYPICAL CASE STUDIES: PRACTICAL APPLICATION OF MACHINE LEARNING IN NETWORK ATTACK DETECTION

5.1 Case Background and Research Object

In recent years, a large Internet company in China has encountered several serious network attacks, which not only led to user data leakage, but also caused business interruption and huge economic losses. In order to cope with increasingly complex and frequent network threats, the company has decided to introduce machine learning technology to enhance its ability to detect network attacks. The research subjects include internal network traffic data, historical attack records, and real-time monitoring data of the company. The network environment is complex, involving multiple protocols and a large number of devices, with a huge amount of data, making it suitable for the application of machine learning algorithms. By analyzing and processing these data, potential attack patterns can be identified, improving the level of network security protection. The company's cybersecurity team selected multiple machine learning algorithms for experimentation, including support vector machines (SVM), decision trees, and random forests, with the aim of finding the most suitable detection method for its network environment and data features.

5.2 Implementation process of machine learning algorithms

During the project implementation process, the collected data was first preprocessed. Data preprocessing includes cleaning, normalization, and dimensionality reduction to remove noise and redundant information, ensuring data quality. Subsequently, feature extraction was carried out to extract key indicators such as traffic characteristics, protocol characteristics, and time characteristics from network traffic data. In order to select the optimal machine learning algorithm, the research team compared and optimized multiple algorithms. The Support Vector Machine (SVM) algorithm maps data to a high-dimensional space through kernel functions, effectively improving classification accuracy; The decision tree algorithm recursively splits the data space to generate easily interpretable classification rules; The random forest algorithm enhances the stability and robustness of detection by constructing multiple decision tree models. In practical applications, the research team inputs network traffic data into a trained machine learning model for real-time detection and analysis. The results show that the random forest algorithm performs the best in processing large-scale and high-dimensional data, with a detection accuracy of 95% and a false positive rate of less than 5%. In addition, the system response time is significantly reduced, allowing for timely detection and handling of potential attacks. This successful case not only demonstrates the potential of machine learning in network attack detection, but also provides valuable experience and reference for other similar projects.

6. CASE EFFECTS AND APPLICATION PROSPECTS

6.1 Improvement of Detection Accuracy and Efficiency

After the introduction of machine learning technology, the Internet company has made significant progress in network attack detection. By applying and optimizing algorithms such as Support Vector Machine (SVM), Decision Tree, and Random Forest, the detection accuracy has been significantly improved. In practical applications, the random forest algorithm performs the best with its ability to handle large-scale and high-dimensional data, increasing detection accuracy to 95%, a significant improvement compared to the traditional method's 80%. In addition, the false alarm rate has significantly decreased from 30% to below 5%, greatly reducing the resource waste and misoperation caused by false alarms. While improving detection accuracy, the response speed of the system has also been greatly improved. Through real-time analysis of network traffic, machine learning algorithms can quickly identify and process potential threats, reducing response time to seconds and ensuring the immediacy and effectiveness of network security protection. In a simulated attack drill targeting the company, the system successfully prevented multiple complex attacks and issued alerts within seconds of the attack, demonstrating its efficiency and reliability in actual combat. This achievement not only enhances the company's network security level, but also provides valuable practical experience and technical reference for other enterprises.

6.2 Future Development Direction of Machine Learning in the Field of Network Security

With the continuous evolution of network attack methods, the application prospects of machine learning in the field of network security are becoming increasingly broad. In the future, the introduction of advanced algorithms such as deep learning and reinforcement learning will further enhance the intelligence level of network attack detection. Deep learning, through a multi-layer neural network structure, can automatically extract and learn complex features from data, adapting to more diverse and covert attack methods. In future applications, deep learning is expected to improve detection accuracy to over 99% and further reduce false alarm rates. Reinforcement learning, through strategy optimization, can self adjust and optimize detection strategies in constantly changing network environments, improving the adaptability and robustness of the system. In a research project of a large technology company in China, preliminary experiments showed that reinforcement learning algorithms have significantly better response capabilities to unknown attacks in simulated environments than traditional methods. In addition, the combination of machine learning with big data analysis, cloud computing, and blockchain technology will drive network security protection into a new stage of intelligence and automation. Through big data analysis, massive network traffic can be monitored and analyzed in real-time to identify potential threats; Cloud computing provides powerful computing capabilities and flexible deployment methods; Blockchain technology ensures the integrity and security of data, providing new solutions for network security. In the future, the application of machine learning in the field of network security will become more profound and extensive, providing solid technical support for addressing increasingly complex network threats.

7. CONCLUSION

After introducing machine learning technology, network attack detection has achieved significant improvements in accuracy and efficiency. By applying and optimizing algorithms such as support vector machine, decision tree, and random forest, the detection accuracy has been significantly improved, the false alarm rate has been significantly reduced, and the response speed has also been greatly improved. These achievements not only enhance the network security level of enterprises, but also provide valuable practical experience and technical references for other similar projects. In the future, with the further development of advanced algorithms such as deep learning and reinforcement learning, the level of intelligence in network attack detection will continue to improve. Combining big data analysis, cloud computing, and blockchain technology, network security protection will enter a new stage of intelligence and automation, providing solid technical support for addressing increasingly complex network threats. Through continuous optimization and innovation, network security will become more stable and provide strong guarantees for the development of the information society.

REFERENCES

[1] Wu, Xiaomin, et al. "Jump-GRS: a multi-phase approach to structured pruning of neural networks for neural decoding." Journal of neural engineering 20.4 (2023): 046020.

[2] Miao, Junfeng, et al. "Secure and Efficient Authentication Protocol for Supply Chain Systems in Artificial Intelligence-based Internet of Things." IEEE Internet of Things Journal (2025).

[3] Li, Huaxu, et al. "Enhancing Intelligent Recruitment With Generative Pretrained Transformer and Hierarchical Graph Neural Networks: Optimizing Resume-Job Matching With Deep Learning and Graph-Based Modeling." Journal of Organizational and End User Computing (JOEUC) 37.1 (2025): 1-24.

[4] Su, Tian, et al. "Anomaly Detection and Risk Early Warning System for Financial Time Series Based on the WaveLST-Trans Model." (2025).

[5] Zhang, Yujun, et al. "MamNet: A Novel Hybrid Model for Time-Series Forecasting and Frequency Pattern Analysis in Network Traffic." arXiv preprint arXiv:2507.00304 (2025).

[6] Peng, Qucheng, Ce Zheng, Zhengming Ding, Pu Wang, and Chen Chen. "Exploiting Aggregation and Segregation of Representations for Domain Adaptive Human Pose Estimation." In 2025 IEEE 19th International Conference on Automatic Face and Gesture Recognition (FG), pp. 1-10. IEEE, 2025.

[7] Wang, Hao. "Joint Training of Propensity Model and Prediction Model via Targeted Learning for Recommendation on Data Missing Not at Random." AAAI 2025 Workshop on Artificial Intelligence with Causal Techniques. 2025.

[8] Pinyoanuntapong, Ekkasit, et al. "Gaitsada: Self-aligned domain adaptation for mmwave gait recognition." 2023 IEEE 20th International Conference on Mobile Ad Hoc and Smart Systems (MASS). IEEE, 2023.

[9] Zheng, Ce, et al. "Diffmesh: A motion-aware diffusion framework for human mesh recovery from videos." 2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). IEEE, 2025.

[10] Wang, Z., Chew, J. J., Wei, X., Hu, K., Yi, S., & Yi, S. (2025). An Empirical Study on the Design and Optimization of an AI-Enhanced Intelligent Financial Risk Control System in the Context of Multinational Supply Chains. Journal of Theory and Practice in Economics and Management, 2(2), 49–62. Retrieved from https://woodyinternational.com/index.php/jtpem/article/view/208

[11] Ding, C.; Wu, C. Self-Supervised Learning for Biomedical Signal Processing: A Systematic Review on ECG and PPG Signals. medRxiv 2024.

[12] Zhang, Yuhan. "InfraMLForge: Developer Tooling for Rapid LLM Development and Scalable Deployment." (2025).

[13] Hu, Xiao. "GenPlayAds: Procedural Playable 3D Ad Creation via Generative Model." (2025).

[14] Li, X., Wang, X., & Lin, Y. (2025). Graph Neural Network Enhanced Sequential Recommendation Method for Cross-Platform Ad Campaign. arXiv preprint arXiv:2507.08959.