Proceedings of CONF-SEML 2025 Symposium: Machine Learning Theory and Applications
DOI: 10.54254/2755-2721/166/2025.TJ23763

AI-enabled Cybersecurity: An In-depth Look at Technologies, Applications, and Challenges

Jinye Xi

Tianjin Chengjian Universit, Tianjing, China
2900017584@qq.com

Abstract: The problem of network security is becoming more and more serious, which brings huge losses to individuals, enterprises and countries. Concurrently, artificial intelligence(AI) has made remarkable progress, and its application in the field of network security has gradually received attention. This paper focuses on the application of AI in the field of network security, and comprehensively compares AI-driven approaches with traditional network security protection methods by using literature research method, case analysis method and comparative analysis method. The objective is to assess the effectiveness and challenges of AI technologies in enhancing network security. The study found that artificial intelligence, relying on machine learning, deep learning and natural language processing and other technologies, performs well in intrusion detection and defense, malicious code identification and defense, and network security situation awareness, which can effectively improve the level of network security protection. However, it also faces many challenges in its application. To address these challenges, the paper proposes methods to enhance the application of AI in network security protection, such as improving the diversity and quality of data, simulating different network environments and attack scenarios, encrypting the collected data, and evaluating models with multiple verification methods.

Keywords: artificial intelligence, Network security, Intrusion detection, Data privacy

1. Introduction

In recent years, with the complex changes in the international situation, cyberspace has become a new battlefield for competition and game among countries. Cyberattack methods have been continuously upgraded, with attacks increasing in both scale and scope. At the individual level, cyber attacks lead to the disclosure of personal privacy, threatening property security and disrupting everyday life. For enterprises, network security incidents may lead to the disclosure of important information such as core business secrets and customer data, resulting in huge losses. At the national level, cyberattacks on infrastructure such as energy, transportation, and finance will pose a serious threat to economic operation, social stability, and national security.

Meanwhile, in the domain of cybersecurity, AI is receiving growing attention for its potential to monitor network traffic in real time and detect anomalous behavior, thereby significantly enhancing cybersecurity defense mechanisms [1]. Currently, the main research fields of re-network security

DOI: 10.54254/2755-2721/166/2025.TJ23763

focus on intrusion detection and defense, malware detection and network security threat prediction [2]. By applying machine learning and deep learning algorithms to analyze data such as network traffic and system logs, various types of network attacks can be accurately identified, such as port scanning, malware transmission, distributed denial-of-service (DDoS) attacks, etc., thereby improving the detection accuracy and efficiency of intrusion detection systems (IDS) and reducing the false positive rate. Moreover, combining the deep learning model to train the characteristics of malware can effectively identify new malware and greatly improve the detection accuracy and speed. In addition, the combination of big data analytics and machine learning technologies can analyze and model historical security data, predict the occurrence of cybersecurity events and formulate preventive strategies in advance.

Despite these advantages, several challenges persist, including limitations in accessing high-quality training data, ensuring data privacy, and enhancing the robustness and reliability of AI models under increasingly sophisticated attack scenarios. Therefore, this paper analyzes the application of artificial intelligence in the field of network security through literature review, case analysis and comparative analysis. By comparing AI technology in network security with traditional network security protection methods, this paper aims to deeply analyze the specific application effects and problems faced by artificial intelligence in network security, and point out the direction for improving and perfecting the application of artificial intelligence technology. With the help of machine learning, deep learning and other algorithms, the application of artificial intelligence in the field of network security provides a new perspective and method for the development of network security theory.

2. Challenges of network security

2.1. Limitations of traditional network security technologies

Traditional IDS and intrusion prevention systems (IPS) mainly rely on predefined rule sets to identify attack behavior [4]. These rules are written based on known attack patterns and characteristics. As a result, they struggle to detect and prevent previously unknown or zero-day attacks—exploits that target undiscovered vulnerabilities without available patches—leaving systems exposed to significant security risks.

In addition, with the continuous expansion of network scale and the increasing complexity of services, the amount of data generated in the network has exploded. Traditional security technology often have performance bottlenecks when processing such a large amount of data, resulting in low detection efficiency. For example, the traditional firewall needs to check and match each packet one by one when facing a large amount of network traffic, which consumes a lot of computing resources and time. When the traffic is too large, the firewall may experience processing delay or even packet loss. As a result, the firewall cannot effectively filter and defend network traffic in a timely manner.

Moreover, traditional security solutions often lack the flexibility and adaptability required to address evolving network environments. For example, in the cloud computing environment, the dynamic migration and elastic expansion of virtual machines blur the network topology and security boundaries, and traditional network security technologies are difficult to effectively protect these dynamically changing resources.

2.2. Emerging trends in cyber attacks

DOI: 10.54254/2755-2721/166/2025.TJ23763

Currently, the network attack means presents the characteristics of diversification, complexity and intelligence, which is specifically reflected in the fact that attackers are no longer limited to a single attack mode, but use a variety of technologies and means to implement attacks. In addition to the usual DDoS attacks, malware spread, phishing and other attack methods, attackers will also use social engineering, supply chain attacks and other new means to break through the network security defense. Social engineering attacks obtain sensitive information of users, such as user names and passwords, through deception and induction, so as to achieve illegal access to the system. Supply chain attacks are attacks on downstream users by attacking the supply chain of software or hardware suppliers and planting malicious code or vulnerabilities in products. The complex trend of cyber attacks makes the process of attack more covert and difficult to prevent. The attacker will use a variety of technical means to confuse the attack behavior and avoid the security detection. In addition, with the development of artificial intelligence technology, cyber attacks have also shown a trend of intelligence. Attackers use artificial intelligence technology to automatically analyze the vulnerabilities of the target system, formulate more accurate attack strategies, and even adjust the attack mode in real time according to the feedback of the defense system.

3. Advantages of AI

3.1. Efficient threat detection capability

Artificial intelligence has demonstrated excellent threat detection capabilities in the field of cybersecurity, mainly due to its characteristics of quickly detecting anomalies and identifying potential threats [5]. Machine learning, for instance, can build a model of normal network behavior by learning from normal network behavior data. When the actual behavior data in the network has a large deviation from the model, the algorithm can quickly detect the abnormal situation. In corporate networks, AI can monitor employee network access behavior in real time, including websites visited, applications used, data transfer volume, and more. Once unusual behavior is detected, alerts can be issued in time to alert possible security threats, such as data breach risk or malware infection. Meanwhile, deep learning models can automatically learn complex patterns and features in network traffic data to accurately identify various types of cyber attacks. In the field of image recognition, convolutional neural networks (CNNS) can be used to analyze the characteristic images of network traffic, and can accurately identify common network attacks such as DDoS attacks and SQL injection attacks by learning the characteristics of normal traffic and attack traffic. Recurrent neural network (RNN), which excel at processing time series data, can be used to analyze the time series changes of network traffic and capture the characteristics of attack behavior in time dimension, so as to discover potential attack threats in time.

3.2. Powerful data analysis and processing capabilities

By establishing appropriate models, machine learning and deep learning algorithms can quickly analyze large-scale network traffic data, including data classification, clustering and prediction [6]. In network traffic analysis, machine learning algorithms can classify network traffic data according to different characteristics, such as protocol type, source IP address, and destination IP address, helping security personnel find potential security problems based on the distribution of network traffic. Deep learning, with its ability to autonomously learn abstract representations from raw data, can extract critical information from network logs, and correlate these attributes to identify potential threats. In addition, through the analysis of historical network security data, AI can learn the

DOI: 10.54254/2755-2721/166/2025.TJ23763

characteristics and rules of different types of attacks, enabling the creation of predictive models. When new data appears, AI can compare it with established attack models to determine if there are similar attacks.

3.3. Automated response and decision making

Automated response and decision making by artificial intelligence in network security greatly improves the efficiency and timeliness of network security protection.

With automated responses, AI automatically initiates traffic cleaning mechanisms when DDoS attacks are detected and automatically quarantines infected hosts when malware infections are detected, based on preset rules and machine learning models. AI can also dynamically adjust security policies of intrusion detection system by analyzing network traffic data and attack characteristics in real time. In addition, AI can collaborate with various cybersecurity systems, including firewalls, antivirus tools, and IPS, enabling coordinated defense strategies. Integration with Security Information and Event Management (SIEM) platforms allows for centralized threat monitoring, event correlation, and situational awareness, ultimately leading to more informed and autonomous security decision-making.

3.4. Data quality and privacy protection

Data quality is crucial for AI models and there are many factors that affect data accuracy [2]. Inaccurate data used for model training will cause the model to misjudge, and then affect the network security protection. In addition, when collecting malware samples or network traffic data, the data is often missing due to channel and scope limitations, making the model unable to learn comprehensive features, and thus reducing the ability to cope with complex security scenarios. In addition, the data required by the network security model contains a large amount of privacy information, which is easy to leak or abuse if it is improperly protected in the process of collection, storage, transmission and use, and brings serious losses to users.

3.5. Robustness and reliability of artificial intelligence model

Robustness against adversarial attacks is a critical requirement for AI applications in cybersecurity. The attacker performs counter-attacks by making small perturbations to the input data, such as tampering with network traffic or malware samples, making the intrusion detection and malware detection models misjudge, thereby bypassing detection. The reliability of the model is also affected by many factors, such as changes in the network environment and new attack means. If the model cannot adapt in time, the detection performance will decline, and there will be missing and false positives.

3.6. Interpretability of AI algorithm

The decision-making process of artificial intelligence algorithms is difficult to explain, especially deep learning algorithms, whose complex structures and mechanisms make the decision-making process opaque. Taking the intrusion detection model based on deep learning as an example, it is difficult for security personnel to know the basis for the model to judge the attack traffic, unable to effectively evaluate its accuracy and reliability, and difficult to optimize the model. Uninterpretable algorithms can also lead to security vulnerabilities. Attackers may exploit blind spots or

DOI: 10.54254/2755-2721/166/2025.TJ23763

inconsistencies within the model's logic—flaws that security personnel may not detect due to the model's opacity.

4. Strategy

4.1. Data processing and privacy protection technology

In terms of data processing and privacy protection technology, data enhancement technology can significantly improve the diversity and quality of data by transforming and expanding the original data, and provide more abundant materials for the training of artificial intelligence models [3]. In the field of network security, augmented traffic data can be generated by simulating various network environments and attack scenarios, thereby improving the model's adaptability to different condition. In the data collection stage, the collected data can be encrypted to ensure the data security. During data transmission, encryption protocols such as SSL/TLS are used to encrypt data and prevent data from being stolen or tampered with. When data is stored, encryption algorithms such as AES (Advanced Encryption Standard) are used to encrypt sensitive data so that only authorized users can decrypt and access the data. In addition, federated learning technology allows multiple participants to train data locally without sharing the original data, and only the parameters or intermediate results of the model are uploaded to a central server for aggregation and updating, thereby effectively protecting data privacy by avoiding the transmission process of the original data.

4.2. Enhancing model robustness and reliability

By introducing adversarial attacks in the training process, the model learns to resist attacks, thus enhancing the robustness of the model [7]. When training an AI-based intrusion detection model, adversarial samples generated by adding tiny perturbations to normal network traffic data are added to the training data to deceive the model into making wrong judgments. Through adversarial training, the model can gradually adapt to adversarial samples during the learning process, and can identify small changes in attack traffic to avoid attackers bypassing detection.

Model fusion is also an effective means to improve model reliability and stability. By integrating multiple models, it is possible to leverage the strengths of different approaches. For example, in malware detection, machine learning-based and deep learning-based malware detection models can be integrated. Machine learning models have strong interpretability and can provide intuitive detection results and interpretation, while deep learning models have strong feature learning capabilities and can handle complex data patterns. The combination of the two can improve the accuracy and reliability of malware detection.

Multiple validation is an important step to ensure the accuracy and reliability of the model. After the model training is completed, a variety of verification methods are used to evaluate the model, including cross-verification and independent test set verification. Cross-validation divides the data set into multiple subsets, comprehensively evaluates the performance of the model through multiple training and validation, and avoids the overfitting phenomenon of the model on a specific data set. Independent test set validation evaluates the model using a test data set that is completely independent of the training set to check the model's generalization ability to unknown data. In the evaluation of intrusion detection model, cross verification and independent test set verification can ensure that the model can accurately detect intrusion behavior in different network environments and attack scenarios.

DOI: 10.54254/2755-2721/166/2025.TJ23763

4.3. Development and application of interpretable artificial intelligence algorithms

In the field of network security, the research and application of interpretable artificial intelligence algorithms are of great significance [8]. Decision tree algorithm, for example, is a relatively intuitive interpretability model, which constructs a tree structure to make decisions [3]. In intrusion detection, the decision tree divides traffic data from the root node according to different characteristics of network traffic, such as source IP address, destination IP address, port number, and protocol type, until the leaf node classifies the traffic as attack traffic [9]. This process allows security analysts to clearly understand how decisions are made and what features the model relies on, fostering greater trust in the system.In addition, by visualizing the decision-making process and feature representation of the model, security personnel can more intuitively understand the working principle and decision basis of the model. For deep learning models, visualization tools can highlight aspects such as feature maps and attention mechanisms. In the CNN-based malware detection model, the visual feature map can present the areas and features that the model pays attention to when processing malware samples, so as to help security personnel understand the process of model extraction and using data features to make decisions.

Local interpretation methods provide local explanations and explanations for specific decision results. The ELocal Interpretable Model - agnostic Explanations (LIME) algorithm can provide local interpretations of the predictions of deep learning models. It explains the decision-making process of the model by generating a set of explanatory samples similar to the original data and analyzing the impact of these samples on the model's prediction results.

5. Conclusion

This study deeply discusses the application of artificial intelligence in the field of network security, and comprehensively analyzes its technical principles, application cases, challenges and countermeasures [10]. Artificial intelligence technologies, including machine learning, deep learning and NLP, have brought new opportunities for network security protection with their powerful data analysis and pattern recognition capabilities .

In the intrusion detection and defense system, the technology based on artificial intelligence can quickly and accurately identify intrusion behaviors through the analysis of network traffic and system logs and other data, and has a high detection accuracy rate and the detection ability of new attacks [8]. Through the comprehensive analysis of massive network data, artificial intelligence can achieve a comprehensive and dynamic assessment of network security conditions and provide strong support for security decisions. However, in terms of technological innovation and optimization, AI needs to enhance the protection of data and privacy, further improve the robustness and reliability of models, and research and application of interpretable AI algorithms.

This paper has shortcomings in the depth of content and research methods. Future studies can focus on the integration of emerging technologies, ensure network security by using machine learning and deep learning, improve the automated security defense capability of the system, and enable the artificial intelligence system to deal with security threats more intelligently, reduce labor costs, and improve protection efficiency and accuracy.

References

[1]Chen B. (2024). Research and Application of Artificial Intelligence-based Network Security Situational Awareness Technology. Telecom World, 2024, 31(10): 31-33. DOI: 10.3969/j.issn.1006-4222.2024.10.011

DOI: 10.54254/2755-2721/166/2025.TJ23763

[2]Ali, S., Wang, J., & Leung, V. C. M. (2025). AI-driven fusion with cybersecurity: Exploring current trends, advanced techniques, future directions, and policy implications for evolving paradigms– A comprehensive review. Information Fusion, 118, 102922.

[3]Ilieva, R., & Stoilova, G. (2024, September 17 - 19). Challenges of AI - Driven Cybersecurity. In 2024 XXXIII International Scientific Conference Electronics (ET) (pp. n/a). IEEE.

[4]Achuthan, K., Ramanathan, S., Srinivas, S., & Raman, R. (2024). Advancing cybersecurity and privacy with artificial intelligence: current trends and future research directions. Frontiers in Big Data, 7, 1497535.

[5]Fu, H. M. & Li, D. (2025). Research on the Application of Artificial Intelligence in Network Security of University Computer Rooms. Scientific and Technological Innovation, (09):105-108.

[6]Kaur, R., Gabrijelčič, D., & Klobučar, T. (2023). Artificial intelligence for cybersecurity: Literature review and future research directions. Information Fusion, 97, 101804.

[7]Zhang, C.T., Cheng, C., Wang, H.L, Bao, A.Y. & Ding, N.Z. (2025). Application of Artificial Intelligence Technology in Computer Network Security Protection. Computer Knowledge and Technology, 21(01): 102-104+107. DOI:10.14004/j.cnki.ckt.2025.0095.

[8]Jada, I., & Mayayise, T. O. (2024). The impact of artificial intelligence on organisational cyber security: An outcome of a systematic literature review. Data and Information Management, 8(2), 100063.

[9]Wang, T.Y. & Guo, Y. (2025). Analysis of Artificial Intelligence Technology in Network Security Defense. Northeast Electric Power Technology, 46(02): 40-42. DOI: 10.3969/j.issn.1004-7913.2025.02.010

[10]Hua,, C., Pu, J.L. & Zhang, T.T. (2024). Application of Artificial Intelligence Technology in Cyberspace Security Defence. Digital Technology and Application, 42(10):60-62.