Artificial Intelligence-driven Cybersecurity Applications and Challenges

Zhaoyue Bai[1], Hui Miao[1*], Junfeng Miao[1], Nan Xiao[2], Xiaoxue Sun[2]

[1] Henan Normal University, Xinxiang, 453000, China

[2] University of Science and Technology Beijing, Beijing, 100083, China

* 2684047661@qq.com

https://doi.org/10.70695/AA1202502A09

Abstract

In the era of big data, novel cyber attack methods are emerging endlessly. Traditional cybersecurity defense measures have become insufficient to counter these new threats, while the rapid development of artificial intelligence (AI) technology provides critical solutions for cybersecurity defense systems. The application of AI in the field of cybersecurity is primarily manifested in the following aspects. Firstly, intelligent threat monitoring and intrusion detection systems enable real-time indentification of anomalous behaviors. Secondly, machine learning-based automated defense and response mechanisms significantly enhance the efficiency of security incident handling. Thirdly, intelligent vulnerability and mining technologies greatly improve the capability to discover security flaws. Lastly, novel authentication technologies such as biometric recognition have reconstructed access control systems. However, the application of AI in cybersecurity still faces numerous challenges. The maturity of the technology needs further improvement and there is a severe shortage of interdisciplinary talent proficient in both AI and cybersecurity. The rapid evolution of attack methods increases defensive difficulties and the resulting ethical issues urgently require resolution. With the continuous advancement and refinement of technology, AI will undoubtedly build a more intelligent and efficient protective barrier for cebersecurity defense systems.

Keywords Artificial Intelligence; Cybersecurity; Security Protection; Threat Detection; Efficient

1 Introduction

With the continuous development of big data, cloud computing and digital transformation, AI has been rapidly applied across multiple domains. Cyberspace is the fifth major strategic competition and a core pillar for socioeconomic development [1-2]. Currently, the increasing sophistication of cyber threats and the growing intelligence of attack methods pose severe challenges to tradition cybersecurity defense systems.

AI refers to the technology and science enabling computer systems or machine models to extend and augment human intelligence, allowing them to perform tasks that typically require human cognitive capabilities, such as learning, reasoning, problem-solving, language comprehension and decision-making [3-4]. In its early stages, AI development was constrained by computational limitations, leading to slow progress. However, in recent years, with technological maturation, AI has achieved qualitative breakthroughs. From Turing's initial exploration of machine intelligence to the creation of AI, AI has evolved dramatically within just a few decades [5]. The three core pillars of AI are alogrithms, data and computing power. Advances in algorithmic and computational capabilities, coupled with the advent of the big data era, have propelled rapid AI development. Currently, prominent AI technologies include deep learning, machine learning, natural language processing, among others [6].

Machine Learning is a branch of AI that enables computer systems to automatically learn patterns and rules from data using algorithms, thereby making predictions and decisions for future problems [7]. As a mainstream subfield of AI, machine learning eliminates the need for traditional hard-coded rules, allowing programs to operate and respond to events without explicit programming. Currently, the two most prevalent learning paradigms are supervised learning and unsupervised learning [8]. Deep Learning, a subset of machine learning, is based on multi-layered neural network architectures and is particularly suited for processing large-scale unstructured data, such as images, videos, and natural language [9]. Its defining characteristic is the ability to automatically extract hierarchical features

through layered neural networks, progressively refining higher-level representations. Core components of DL include neural network such as Convolution Neural Networks(CNN), Recurrent Neural Network(RNN), Transformer and Generative Adversarial Networks(GAN). In recent years, deep learning has achieved breakthroughs in image recognition, speech processing and related domains [10]. Nature Language Processing(NLP) focuses on enabling computers to understand and interact with language. Its technical essence lies in establishing a mapping between machine-interpretable symbolic systems and the semantic connotation of human language. Key applications include intelligent customer service, sentiment analysis, machine translation、text generation and semantic understanding [11].

Since the inception of computer networks in the 1950s to the establishment of today's vast global internet village, the internet has interconnected people worldwide, giving rise to cybersecurity issues. Currently, cybersecurity faces diverse and complex threats, including malware attacks, phishing attacks, Advanced Persistent Threats (APT), Distributed Denial of Service (DDoS) attacks and insider threats.

Malware includes viruses, worms, trojans, ransomware, etc. [12]. Such attacks aim to steal money through data destruction, information theft, extortion, or disabling systems. For example, the notorious ransomware wannacry exploited windows vulnerabilities and caused billions in economic losses during its global outbreak on May 12, 2017. Phishing attacks deceive users into clicking malicious links, with objectives similar to malware. APT is highly covert, long-term and damaging cyberattacks [13]. DDoS attacks involve multiple attackers targeting one or more systems in a coordinated, large-scale assault, often involving hundreds or even thousands of compromised hosts [14]. These attacks typically paralyze target systems, leading to significant financial losses.

Beyond external threats, insider threats are also a critical cybersecurity concern. Personnel within organizations may inadvertently or maliciously leak sensitive information due to insufficient security awareness or personal gain [15], resulting in system damage or operational disruptions. Compared to external threats, insider threats are harder to detect, as attackers often possess legitimate access permissions. With cyberattacks becoming increasingly sophisticated and automated, traditional security defenses struggle to counter emerging threats like APT and zero-day attacks. Research into AI-driven cybersecurity technologies can enhance defensive capabilities and foster a more secure cyberspace.

2 Applications of AI in Cybersecurity

2.1 Threat Detection and Intrusion Identification

Threat detection and intrusion identification are core components of cybersecurity, aiming to discover, analyze and block malicious activities in real time to protect user systems from attacks. For emerging attack methods, the following examples provide countermeasures.

Malware Detection and Analysis System(MDS)

The MDS system, based on AI, is designed for real-time monitoring, identification and analysis of cybersecurity threats [16]. Leveraging deep learning and NLP technologies, MDS can extract critical information from vast amounts of threat intelligence and rapidly classify malware. At the data collection layer, it gathers logs from firewalls, servers, endpoints and other devices, capturing network traffic via SPAN ports or network taps [17]. It employs Security Orchestration, Automation and Response to automatically handle low-risk alerts and collaborates with firewalls, Endpoint Detection and Response and other systems to block attacks, enabling automated response. Compared to traditional Intrusion Detection/Prevention Systems, MDS excels by integrating AI, allowing it to detect complex cyber threats, achieve full-stack monitoring and provide deeper detection dimensions.

Network Traffic Analysis

With the rapid evolution of malicious traffic attack techniques, network traffic analysis has become an effective defensive measure. It primarily encompasses two aspects: network traffic classification and network traffic prediction [18].The former is currently a major research focus, with studies covering: supervised machine learning classification methods, unsupervised machine learning classification methods, semi-supervised machine learning classification methods and neural network-based network traffic classification methods. The latter is an approach combining data and algorithms to predict security risks before they materialize. Recent research predominantly leverages traditional machine learning algorithms and deep learning techniques such as neural networks.

Network Anomaly Detection

Network anomaly detection is a system built on traditional machine learning that conducts real-time analysis of network traffic and user behavior to identify potential cyber threats. Numerous related

studies exist in this field. Examples include Time-series anomaly detection methods based on Kolmogorov-Arnold representation theory and unsupervised adversarial model-based network anomaly detection algorithms. By identifying anomalous behavior, network anomaly detection systems can automatically initiate countermeasures, analyzing data autonomously to uncover potential risks [19]. Moreover, these systems possess self-learning and optimization capabilities, continuously improving detection accuracy while maintaining the ability to detect previously unknown attack vectors.

2.2 Security Defense and Response

Situational Awareness

Cybersecurity situational awareness refers to a comprehensive security capability that involves real-time monitoring, analysis and assessment of various security elements within a network system, dynamically grasping the overall security status and predicting potential threats [20]. In data collection and analysis, deep learning algorithms demonstrate significant advantages. Convolutional Neural Networks (CNN) can effectively process spatial features in network traffic, while Recurrent Neural Networks (RNN) and Long Short-Term Memory networks (LSTM) have achieved breakthroughs in log analysis, traffic monitoring and vulnerability feature extraction [21]. In attack attribution, deep learning-based models can effectively trace attackers' covert paths and even accurately locate attack sources through behavioral pattern recognition, greatly improving incident response efficiency [22].

Active Defense

Modern cybersecurity threats have become increasingly complex and diverse, rendering traditional manual monitoring and defense methods inadequate. The application of AI in cybersecurity has enabled a shift from passive to active defense [23]. Using machine learning and deep learning algorithms, AI can continuously detect and identify abnormal behavioral patterns in networks, such as malware, DDoS attacks and phishing sites, automatically recognizing and intercepting these threats upon detection. Machine learning-based AI algorithms continuously learn from vast datasets, predicting potential risk areas and automating defensive tasks [24]. Compared to human analysts, AI excels at big data analysis, enabling cybersecurity defense systems to achieve faster and more efficient responses.

2.3 Vulnerability Exploitation Technology

Vulnerabilities refer to computational logic errors discovered in software and certain hardware components (e.g., firmware). These vulnerabilities can be exploited by hackers to gain access to systems or networks and execute malicious actions. Research on AI-based software vulnerability mining technologies facilitates the rapid discovery and timely remediation of system security flaws. Software vulnerability mining involves analyzing the source code or executable code of software to detect exploitable defects.

Static and Dynamic Analysis-Based Vulnerability Mining

Static analysis-based modeling is categorized into graph-based static analysis and static modeling with data modeling. The former models program characteristics as control flow graphs, data flow graphs or other graphical representations to identify errors in the program. The latter employs lattice theory-based abstraction methods for computation. Dynamic analysis, on the other hand, monitors the runtime behavior of a given program by analyzing specific input data. However, this approach cannot analyze the entire program and may overlook vulnerabilities [25].

Intelligent Fuzzing-Based Vulnerability Mining

Fuzzing is a widely used automated black-box testing method. The basic process involves target identification, input identification, data construction, anomaly detection and anomaly verification. Researchers supply the target program with a large volume of specially crafted or random inputs and monitor for anomalies to uncover vulnerabilities [26]. Since this is a black-box testing method, the results are highly unstable and prolonged efforts may yield no findings. Researchers actively refine techniques to improve stability, which have contributed to advancements in fuzzing.

Binary Diffing-Based Vulnerability Mining

Identifying binary executable programs is fundamental to vulnerability mining. The core concept involves comparing pre-patch and post-patch binary files or different compiled versions of the same file to pinpoint critical changes that may introduce risks. This method does not rely on source code and can detect 1-day and n-day vulnerabilities [27].

2.4 Identity Authentication and Access Control

Traditional identity authentication methods, such as static password authentication, dynamic password authentication and security question verification, are no longer sufficient to defend against emerging cyber threats. AI-based novel identity authentication methods have brought revolutionary changes to user privacy protection [28]. Leveraging natural language processing (NLP) technology, AI can automatically process and classify security incident reports, assisting security experts in rapidly formulating solutions. Multi-factor authentication (MFA) technology is being adopted to replace traditional authentication methods, integrating biometrics (e.g., facial recognition, voice recognition, iris recognition), behavioral characteristics and cryptography to enhance user security access control [29]. The system can further utilize machine learning and deep learning algorithms to continuously analyze user behavioral patterns to construct models, enabling real-time continuous authentication. If user behavior deviates from the established model, secondary authentication is triggered. Additionally, context-aware technology and real-time threat intelligence dynamically adjust access permissions, forming an adaptive zero-trust security architecture. Finally, LSTM based on deep learning can analyze sequential behavioral patterns, while Generative Adversarial Networks generate adversarial samples to detect spoofing attacks, ensuring robust system security [30].

3 Challenges of AI in Cybersecurity

3.1 Privacy and Ethical Challenges

While AI offers significant advantages in cybersecurity, it also presents inherent risks and challenges. In terms of privacy, excessive data collection may infringe upon user privacy, potentially leading to data misuse and privacy breaches [31]. Empirical studies indicate that 78% of enterprise AI-based security systems currently exhibit issues related to excessive data gathering. On the ethical front, challenges include algorithmic bias, the misuse of surveillance under the pretext of security and ambiguous accountability in cybersecurity incidents. Addressing these issues will require more stringent regulatory frameworks and legal standards in the future.

3.2 Technical Limitation Challenges

The technical constraints of AI in cybersecurity are primarily manifested in four aspects. First, data dependency restricts model performance due to limitations in data quality and coverage, resulting in insufficient generalization capabilities when encountering novel attacks. Second, algorithmic vulnerability renders AI systems susceptible to adversarial attacks, with success rates exceeding 85% in some cases. Third, real-time processing bottlenecks hinder the application of complex models in high-speed network environments, making millisecond-level breakthroughs challenging [32]. Finally, the lack of interpretability impedes deployment in critical decision-making scenarios, as the black-box nature of AI prevents security professionals from understanding its decision logic. These limitations lead to false negatives in AI-powered security systems when confronting emerging attack methods, thereby introducing security risks.

3.3 Practical Application Challenges

Despite the achievements in the new paradigm of integrating AI with cybersecurity, several challenges persist in practical applications. First, obtaining high-quality security data remains difficult, as enterprises are reluctant to share sensitive data, leading to insufficient model training. Second, the high computation costs, particularly the demand for high-performance GPU clusters, pose a significant financial burden for small and medium-sized enterprises. Third, there is a shortage of interdisciplinary professionals proficient in both AI and cybersecurity. Finally, achieving a fully integrated system is challenging due to compatibility issues between traditional security infrastructure and AI-driven systems. Addressing these challenges will facilitate the development of an intelligent and robust cybersecurity defense framework.

While cybersecurity defense mechanisms continue to advance, offensive techniques are simultaneously undergoing significant sophistication. Contemporary attack vectors now incorporate machine learning to generate polymorphic malicious code variants capable of evading detection systems.

Furthermore, social engineering attacks leveraging deep learning can produce highly convincing disinformation, effectively deceiving users into interacting with malicious links. Modern botnets exhibit fundamental architectural differences from traditional variants - AI-enabled network clusters now possess autonomous coordination capabilities. These systems can execute targeted attacks without centralized command structures while simultaneously degrading victim response times [34]. Attackers are increasingly adopting long-term persistence strategies that evade conventional monitoring systems. The continuous evolution of these offensive methodologies presents formidable challenges to existing cyber defense paradigms. The particular concern is the growing capability gap between adaptive attack techniques and static defense systems, necessitating fundamental advancements in protective architectures.

4 Future Development Trends of AI-Based Cybersecurity

4.1 Technological Convergence and Development

The future integration of AI with cybersecurity will extend beyond mere technical enhancements to facilitate a comprehensive reconstruction of security systems. Recent breakthroughs in unsupervised and semi-supervised learning have enabled AI systems to autonomously detect potential threats without extensive human supervision. Concurrently, AI-optimized quantum key distribution protocols are significantly improving the efficiency of quantum-secure communications, presenting novel solutions to emerging quantum security challenges. Furthermore, innovative applications of multimodal AI technology are integrating diverse data sources-including system logs, video surveillance and linguistic data. This technological evolution is further enhanced through interdisciplinary convergence, as cybersecurity progressively incorporates insights from legal studies, psychology and social engineering to elevate its intelligent capabilities [35]. Such multidimensional advancements are collectively driving the transformation from conventional security frameworks to next-generation AI-powered defense ecosystems.

4.2 Autonomous Decision-Making Security Systems

Dynamic Defense

Cybersecurity faces continuously evolving threats, with attackers employing increasingly sophisticated and stealthy techniques. Traditional static defense mechanisms, such as firewalls, have become inadequate against these advanced attacks. AI-driven cybersecurity defense systems demonstrate autonomous learning capabilities, automatically analyzing and processing massive datasets to recognize attack patterns. These systems continuously detect and predict advanced threats while playing a crucial role in mitigating unknown risks [36].

Autonomous Strategy Formulation and Decision-Making

AI-powered cybersecurity defense systems can autonomously develop countermeasures against novel cyberattacks. They perform real-time monitoring of network anomalies and dynamically adjust defense strategies without human intervention, enabling automatic risk identification and self-resolution [37].

4.3 Privacy Computing

The advancement of AI in cybersecurity necessitates the collection, analysis and training of massive datasets. Machine learning, as a critical branch of AI, utilizes collected raw data to train scenario-specific models that mimic human behavior. Federated learning, a distributed machine learning approach, enables decentralized training of partial models on local data, which are then aggregated on a central server to construct a global model. However, attackers can potentially reconstruct original user data by intercepting partial data during decentralized training, making privacy protection during data training particularly crucial [38].Common privacy-preserving techniques include secure multi-party computation and homomorphic encryption.

Privacy-Preserving Aggregation Mechanism Based on Secure Multi-Party Computation (SMPC)

SMPC is a cryptographic protocol that enables multiple mutually distrustful parties to jointly compute a predetermined function and obtain the final result without disclosing their respective private data. The key cryptographic technologies employed include secret sharing and homomorphic encryption [39].By

utilizing SMPC protocols, user data privacy can be effectively protected, facilitating more secure construction of cybersecurity defense systems.

Privacy-Preserving Aggregation Mechanism Based on Homomorphic Encryption

Homomorphic encryption, a cryptographic algorithm, allows untrusted third parties to perform computations on encrypted data without accessing the original plaintext. Based on computational capabilities, it can be classified into three categories: partially homomorphic encryption, somewhat homomorphic encryption and fully homomorphic encryption [40]. Privacy-preserving mechanisms based on this technology will effectively safeguard data privacy, enabling continuous development of cybersecurity technologies while maintaining complete data protection.

5 Conclusion

In the era of rapid AI development, the cybersecurity field is undergoing profound transformation and innovation. This paper systematically examines the innovative applications of AI in cybersecurity, current challenges and future development trends. As a global digital nexus, cyberspace breaks information barriers and transcends geographical limitations while inevitably facing various cyber threats. To address these challenges, researchers have developed multi-layered AI-based defense systems that not only enable real-time monitoring and intelligent identification of cyber threats but also shift security defenses forward through predictive security analytics. These technologies effectively safeguard personal privacy, financial transactions and enterprise data security. The deep integration of AI and cybersecurity is redefining protection paradigms in cyberspace, driving the evolution of cybersecurity technologies from passive defense to proactive intelligent defense systems.

Acknowledgement

This work was supported without any funding.

Conflicts of Interest

The authors declare no conflicts of interest.

References

1. Gursoy, D., & Cai, R. (2025). Artificial intelligence: an overview of research trends and future directions. International journal of contemporary hospitality management, 37(1), 1-17.
2. Grzybowski, A., Pawlikowska–Łagód, K., & Lambert, W. C. (2024). A history of artificial intelligence. Clinics in Dermatology, 42(3), 221-229.
3. Varghese, C., Harrison, E. M., O'Grady, G., & Topol, E. J. (2024). Artificial intelligence in surgery. Nature medicine, 30(5), 1257-1268.
4. Grzybowski, A., Pawlikowska–Łagód, K., & Lambert, W. C. (2024). A history of artificial intelligence. Clinics in Dermatology, 42(3), 221-229.
5. Giudici, P., Centurelli, M., & Turchetta, S. (2024). Artificial Intelligence risk measurement. Expert Systems with Applications, 235, 121220.
6. Sheth, A., Roy, K., & Gaur, M. (2023). Neurosymbolic artificial intelligence (why, what, and how). IEEE Intelligent Systems, 38(3), 56-62.
7. Goar, V., & Yadav, N. S. (2024). Foundations of machine learning. In Intelligent Optimization Techniques for Business Analytics (pp. 25-48). IGI Global.
8. Räz, T., & Beisbart, C. (2024). The importance of understanding deep learning. Erkenntnis, 89(5), 1823-1840.
9. Weng, Y., & Wu, J. (2024). Leveraging artificial intelligence to enhance data security and combat cyber attacks. Journal of Artificial Intelligence General science (JAIGS) ISSN: 3006-4023, 5(1), 392-399.
10. Tulli, S. K. C. (2024). Artificial intelligence, machine learning and deep learning in advanced robotics, a review. International Journal of Acta Informatica, 3(1), 35-58.
11. Chowdhary, K., & Chowdhary, K. R. (2020). Natural language processing. Fundamentals of artificial intelligence, 603-649.
12. Kramer, S., & Bradfield, J. C. (2010). A general definition of malware. Journal in computer virology, 6, 105-114.

13. Connor, G., & Korajczyk, R. A. (1988). Risk and return in an equilibrium APT: Application of a new test methodology. Journal of financial economics, 21(2), 255-289.
14. Douligeris, C., & Mitrokotsa, A. (2004). DDoS attacks and defense mechanisms: classification and state-of-the-art. Computer networks, 44(5), 643-666.
15. Thakur, K., Qiu, M., Gai, K., & Ali, M. L. (2015, November). An investigation on cyber security threats and security models. In 2015 IEEE 2nd international conference on cyber security and cloud computing (pp. 307-311). IEEE.
16. Aslan, Ö. A., & Samet, R. (2020). A comprehensive review on malware detection approaches. IEEE access, 8, 6249-6271.
17. Bensaoud, A., Kalita, J., & Bensaoud, M. (2024). A survey of malware detection using deep learning. Machine Learning With Applications, 16, 100546.
18. Shen, M., Ye, K., Liu, X., Zhu, L., Kang, J., Yu, S., ... & Xu, K. (2022). Machine learning-powered encrypted network traffic analysis: A comprehensive survey. IEEE Communications Surveys & Tutorials, 25(1), 791-824.
19. Hooshmand, M. K., & Hosahalli, D. (2022). Network anomaly detection using deep learning techniques. CAAI Transactions on Intelligence Technology, 7(2), 228-243.
20. Munir, A., Aved, A., & Blasch, E. (2022). Situational awareness: techniques, challenges, and prospects. AI, 3(1), 55-77.
21. Herrmann, L., & Kollmannsberger, S. (2024). Deep learning in computational mechanics: a review. Computational Mechanics, 74(2), 281-331..
22. Mei, Y., Han, W., Li, S., Lin, K., Tian, Z., & Li, S. (2024). A Novel Network Forensic Framework for Advanced Persistent Threat Attack Attribution Through Deep Learning. IEEE Transactions on Intelligent Transportation Systems.
23. Sun, N., Ding, M., Jiang, J., Xu, W., Mo, X., Tai, Y., & Zhang, J. (2023). Cyber threat intelligence mining for proactive cybersecurity defense: A survey and new perspectives. IEEE Communications Surveys & Tutorials, 25(3), 1748-1774.
24. Okoli, U. I., Obi, O. C., Adewusi, A. O., & Abrahams, T. O. (2024). Machine learning in cybersecurity: A review of threat detection and defense mechanisms. World Journal of Advanced Research and Reviews, 21(1), 2286-2295.
25. Damodaran, A., Troia, F. D., Visaggio, C. A., Austin, T. H., & Stamp, M. (2017). A comparison of static, dynamic, and hybrid analysis for malware detection. Journal of Computer Virology and Hacking Techniques, 13, 1-12.
26. Guo, W., Fang, Y., Huang, C., Ou, H., Lin, C., & Guo, Y. (2022). HyVulDect: a hybrid semantic vulnerability mining system based on graph neural network. Computers & Security, 121, 102823.
27. Yang, S., Xu, Z., **ao, Y., Lang, Z., Tang, W., Liu, Y., ... & Sun, L. (2023). Towards practical binary code similarity detection: Vulnerability verification via patch semantic analysis. ACM Transactions on Software Engineering and Methodology, 32(6), 1-29.
28. Albalawi, S., Alshahrani, L., Albalawi, N., Kilabi, R., & Alhakamy, A. A. (2022). A comprehensive overview on biometric authentication systems using artificial intelligence techniques. International Journal of Advanced Computer Science and Applications, 13(4), 1-11.
29. Olabanji, S. O., Olaniyi, O. O., Adigwe, C. S., Okunleye, O. J., & Oladoyinbo, T. O. (2024). AI for Identity and Access Management (IAM) in the cloud: Exploring the potential of artificial intelligence to improve user authentication, authorization, and access control within cloud-based systems. Authorization, and Access Control within Cloud-Based Systems (January 25, 2024).
30. Phiri, J., Zhao, T. J., Zhu, C. H., & Mbale, J. (2011). Using artificial intelligence techniques to implement a multifactor authentication system. International Journal of Computational Intelligence Systems, 4(4), 420-430.
31. Solove, D. J. (2025). Artificial intelligence and privacy. Fla. L. Rev., 77, 1.
32. Dilmaghani, S., Brust, M. R., Danoy, G., Cassagnes, N., Pecero, J., & Bouvry, P. (2019, December). Privacy and security of big data in AI systems: A research and standards perspective. In 2019 IEEE international conference on big data (big data) (pp. 5737-5743). IEEE.
33. Sontan, A. D., & Samuel, S. V. (2024). The intersection of Artificial Intelligence and cybersecurity: Challenges and opportunities. World Journal of Advanced Research and Reviews, 21(2), 1720-1736.
34. Awadallah, A., Eledlebi, K., Zemerly, J., Puthal, D., Damiani, E., Taha, K., ... & Yeun, C. Y. (2024). Artificial intelligence-based cybersecurity for the metaverse: research challenges and opportunities. IEEE Communications Surveys & Tutorials.
35. Kaur, R., Gabrijelčič, D., & Klobučar, T. (2023). Artificial intelligence for cybersecurity: Literature review and future research directions. Information Fusion, 97, 101804.
36. Zheng, Y., Li, Z., Xu, X., & Zhao, Q. (2022). Dynamic defenses in cyber security: Techniques, methods and challenges. Digital Communications and Networks, 8(4), 422-435.
37. Chitimoju, S. (2023). Ethical Challenges of AI in Cybersecurity: Bias, Privacy, and Autonomous Decision-Making. Journal of Computational Innovation, 3(1).
38. Wen, J., Zhang, Z., Lan, Y., Cui, Z., Cai, J., & Zhang, W. (2023). A survey on federated learning: challenges and applications. International Journal of Machine Learning and Cybernetics, 14(2), 513-535.

39. Zhou, I., Tofigh, F., Piccardi, M., Abolhasan, M., Franklin, D., & Lipman, J. (2024). Secure multi-party computation for machine learning: A survey. IEEE Access.
40. Acar, A., Aksu, H., Uluagac, A. S., & Conti, M. (2018). A survey on homomorphic encryption schemes: Theory and implementation. ACM Computing Surveys (Csur), 51(4), 1-35.

Biographies

1. **Zhaoyue Bai** currently studies at Henan Normal University. She is interested in artificial intelligence and cybersecurity.
2. **Hui Miao** currently studies at Henan Normal University. She is interested in artificial intelligence and cybersecurity.
3. **Junfeng Miao** is at Henan Normal University. He is interested in artificial intelligence and cybersecurity.
4. **Nan Xiao** currently studies at University of Science and Technology Beijing. He is interested in artificial intelligence and cybersecurity.
5. **Xiaoxue Sun** currently studies at University of Science and Technology Beijing. She is interested in artificial intelligence and cybersecurity.

人工智能驅動的網絡安全應用與挑戰

白朝悅[1]，苗會[1]，苗俊峰[1]，肖楠[2]，孫曉雪[2]

[1] 河南師範大學，新鄉，中國，453000

[2] 北京科技大學，北京，中國，100083

摘要：在大數據時代，新的網絡攻擊方法層出不窮。傳統的網絡安全防禦措施已不足以應對這些新威脅，而人工智能（AI）技術的快速發展為網絡安全防禦系統提供了關鍵的解決方案。人工智能在網絡安全領域的應用主要體現在以下幾個方面。首先，智能威脅監測和入侵檢測系統能夠實時識別異常行為。其次，基於機器學習的自動化防禦和響應機製顯著提高了安全事件處理的效率。再次，智能漏洞挖掘技術大大增強了發現安全缺陷的能力。最後，生物特征識別等新型認證技術重建了訪問控製系統。然而，人工智能在網絡安全中的應用仍面臨諸多挑戰。技術的成熟度有待進一步提高，同時既精通人工智能又熟悉網絡安全的跨學科人才嚴重短缺。攻擊方法的快速演變增加了防禦難度，由此產生的倫理問題也亟待解決。隨著技術的不斷進步和完善，人工智能無疑將為網絡安全防禦系統構建起更智能、更高效的防護屏障。

關鍵詞：人工智能；網絡安全；安全防護；威脅檢測；高效

1. 白朝悅，目前在河南師範大學學習。她對人工智能和網絡安全感興趣；
2. 苗會，目前在河南師範大學學習。她對人工智能和網絡安全感興趣；
3. 苗俊鋒，目前在河南師範大學。他對人工智能和網絡安全感興趣；
4. 肖楠，目前在北京科技大學學習。他對人工智能和網絡安全感興趣；
5. 孫曉雪，目前在北京科技大學學習。她對人工智能和網絡安全感興趣。