sciendo International Conference KNOWLEDGE-BASED ORGANIZATION
Vol. XXIX No 3 2023

A REVIEW OF ENHANCING INTRUSION DETECTION SYSTEMS FOR CYBERSECURITY USING ARTIFICIAL INTELLIGENCE (AI)

Michal MARKEVYCH, Maurice DAWSON

Illinois Institute of Technology, Chicago, USA
mmarkevych@hawk.iit.edu

Abstract: *The escalating complexity of cyber attacks demands innovative intrusion detection systems (IDS) to safeguard critical assets and data. The study aims to explore the potential of Artificial Intelligence (AI) in enhancing the IDS's ability to identify and classify network traffic and detect anomalous behavior. The paper offers a concise overview of IDS and AI and examines the existing literature on the subject, highlighting the significance of integrating advanced language models for cybersecurity enhancement. The research outlines the methodology employed to assess the efficacy of AI within IDS. Furthermore, the study considers key performance metrics such as detection accuracy, false positive rate, and response time to ensure a comprehensive evaluation. Findings indicate that AI is a valuable asset in enhancing the accuracy of AI for detecting and responding to cyber attacks. Nonetheless, the study also brings to light certain limitations and challenges associated with incorporating AI into IDS, such as computational complexity and potential biases in training data. This research emphasizes the potential of advanced language models like ChatGPT in augmenting cybersecurity solutions and offers insights into overcoming associated challenges for a more robust and effective defense against sophisticated cyber attacks.*

Keywords: ChatGPT, Intrusion Detection Systems, Cybersecurity, AI

1. Introduction

Intrusion Detection Systems (IDS) serve as vital security safeguards, shielding network infrastructures from cyber attacks by identifying unauthorized access and harmful actions. Since their emergence in the mid-80s, they have undergone substantial advancements to stay on par with the growing complexity of computer-related crimes [1]. IDS can be categorized into network intrusion detection (NIDS) and prevention (IPS) systems, which analyze network traffic for signs of malicious activity using signature and statistical anomaly detection as well as heuristic behavioral analysis [2]. Such systems have the capability to identify and potentially avert attacks and malicious actions that conventional security measures like firewalls may miss [3].

The need for improved accuracy in detecting and responding to cyber attacks is paramount due to the growing number of advanced threats that can compromise the confidentiality, integrity, and availability of network systems. According to Pietraszek's estimation, nearly 99% of the intrusion detection alerts are unrelated to cybersecurity concerns, as there are only slight discrepancies observed between regular and malevolent activities [4]. Researchers have proposed various techniques to enhance IDS capabilities, such as using fuzzy logic [1], neural networks (NNs), and support vector machines (SVMs) [2]. These approaches have shown promise in reducing false

DOI: 10.2478/kbo-2023-0072

positives and improving detection rates for different types of attacks, including Distributed Denial of Service (DDoS) attacks [2].

The potential of ChatGPT or similar AI models to enhance IDS capabilities is an area of interest as these models can leverage natural language processing and machine learning techniques to understand complex patterns and behaviors in network traffic. By integrating AI models into IDS, it may be possible to improve the detection of sophisticated attacks, reduce false positives, and enable more efficient response mechanisms. By using ChatGPT or similar AI models, IDS can improve the detection of sophisticated attacks, reduce false positives, and enable more efficient response mechanisms. This paper explores the potential of ChatGPT to improve the accuracy of IDS and enhance their capabilities for cybersecurity.

2. Background

IDS have been an essential component of cybersecurity since the late 1980s. Since then, the field has evolved rapidly in response to the growing complexity and variety of cyber threats. Early intrusion detection systems were primarily focused on securing large, centralized mainframe systems; however, as computer networks became more widespread, IDS expanded to protect these increasingly interconnected systems.

2.1 Traditional Methods of Intrusion Detection Systems

Traditional IDS consist of signature-based detection (SD), anomaly based detection (AD) [5]. Signature Detection (SD) is a method of identifying patterns or sequences in network traffic that match pre-identified attack signatures. This technique is highly effective in detecting known attacks and results in a low false positive rate for such incidents [5]. However, it may not be able to detect emerging or unknown threats, which is a limitation of this approach. Moreover, the signature database needs to be updated continuously to ensure the SD system's efficiency.

Anomaly-based Detection (AD) is a technique that observes network traffic for deviations from regular behavior, which could indicate a potential attack [5]. AD employs machine learning algorithms, statistical analysis, or other methods to establish a standard baseline of normal behavior and identify anomalies. This approach has the ability to detect unknown or novel attacks and is adaptable to evolving network behavior. However, AD has a higher false positive rate compared to Signature Detection (SD), and it requires a training period to establish the baseline of normal behavior. A hybrid approach can be used to address the high false positive and low false negative rates associated with AD.

2.2 Limitations of Traditional Methods in the Face of Evolving Cyber Threats

Traditional intrusion detection methods have served as the backbone of cybersecurity for decades, but their effectiveness has diminished in the face of rapidly evolving and increasingly sophisticated cyber threats such as AI generated attacks [12]. Signature-based detection relies on a database of known threats, which requires constant updates to remain effective. As new threats emerge, traditional IDS may struggle to keep pace, leaving systems vulnerable to novel attacks. Anomaly-based detection methods are prone to false positives, as benign activities that deviate from the norm can trigger alerts. This can lead to an overwhelming number of alerts, which can distract security personnel and reduce overall efficiency. Traditional methods also face scalability issues. As networks and systems grow in size and complexity, traditional IDS may struggle to scale and maintain performance. This can result in slower detection and response times, which can be exploited by attackers.

3. Overview of AI-based IDS

Intrusion Detection Systems (IDS) are essential components of modern network security infrastructure, designed to detect and prevent unauthorized access, misuse, and attacks on computer systems and networks [13]. Traditional IDS rely on signature-based and rule-based methods to detect known threats. However, as the cyber threat landscape evolves, it is becoming increasingly difficult for these traditional approaches to keep up with the rapid proliferation of sophisticated and novel attack techniques [14]. Artificial Intelligence (AI)-based IDS, which leverage machine learning and other AI techniques, have emerged as a promising solution to address these challenges, offering significant advantages over traditional methods in terms of adaptability, pattern recognition, and real-time detection and response capabilities [13].

3.1 Advantages of AI-based IDS over Traditional Methods

One of the key advantages of AI-based IDS is their inherent adaptability. While traditional IDS rely on a fixed set of signatures and rules to detect known threats, AI-based IDS can learn and adapt to new threats and the changing network behavior over time. This enables them to detect previously unseen attacks and anomalies, offering a more robust and proactive defense against ever-evolving cyber threats.

3.2 Pattern Recognition

Another advantage of AI-based IDS is their ability to recognize patterns in large volumes of network data. By using machine learning algorithms, these systems can effectively identify patterns indicative of malicious activity, even when the specific attack vector or method is unknown. This allows AI-based IDS to detect a wide range of threats, including zero-day attacks and advanced persistent threats (APTs), which often go undetected by traditional signature-based IDS [15].

3.3 Real-time Detection and Response

AI-based IDS also excel in real-time detection and response capabilities. Through the use of advanced algorithms and efficient data processing techniques, AI-based IDS can analyze network traffic and detect malicious activity in real-time, allowing organizations to respond to potential security incidents more rapidly and effectively [15]. This significantly reduces the window of opportunity for attackers and minimizes the potential impact of security breaches.

3.4 Challenges and Limitations of AI-based IDS

Despite their numerous advantages, AI-based IDS are not without challenges and limitations [8]. One significant issue is the occurrence of false positives and false negatives, which can lead to an increased workload for security analysts and potential gaps in security coverage [9]. While AI-based IDS are designed to improve detection accuracy, it is essential to continuously refine their algorithms and fine-tune system parameters to minimize these errors.

The computational complexity of AI-based IDS can also pose challenges, particularly for organizations with limited resources [10]. Machine learning algorithms and other AI techniques often require substantial computational power and memory, which may necessitate the deployment of specialized hardware and infrastructure. As such, the cost and resource implications of implementing AI-based IDS must be carefully considered.

Finally, using AI-based IDS raises data privacy concerns, as these systems typically rely on analyzing large volumes of sensitive network data [11]. Ensuring the privacy and security of this data is critical, and organizations must carefully evaluate the potential risks associated with implementing AI-based IDS, including data storage, transmission, and processing practices. Compliance with relevant data protection regulations and implementing appropriate security measures are essential to mitigating these risks.

4. Literature Review

MIT is also working on methods to use machine learning to defend against cyber attacks. Their paper "AI2: Training a big data machine to defend" presents a new method. Their system has four components. A big data processing system, an outlier detection engine, a mechanism to obtain feedback from security analysts, and a supervised learning module. Their system tries to combine the expertise of security experts, and the speed and ability to detect new attacks of machine learning. More specifically, they use unsupervised machine learning. They preferred unsupervised machine learning since labeled data is rare and attacks constantly evolve. In the system, they generate their labels and use a supervised learning algorithm with these labels. The big data processing system is a system that can extract features of different entities from raw data [7]. The outlier detection engine is a system that uses unsupervised learning. It uses the features found in the big data processing system. They use three methods: density, matrix decomposition, or replicator neural networks. The output of this unsupervised system is processed and shown to a security analyst. The security analyst can verify or refute the output. The feedback is fed to a supervised learning algorithm. The supervised learning algorithm learns a model that can use this feedback to predict better whether any new event is normal or abnormal. With more feedback, the system becomes more and more correct.

5. Role of ChatGPT and Similar AI Models in Enhancing IDS

AI-based anomaly detection involves using machine learning techniques, such as unsupervised learning or semi-supervised learning, to identify unusual patterns in network traffic data. Unsupervised learning algorithms, like clustering, can be employed to group similar data points together, thus allowing the AI model to distinguish between normal and abnormal behavior. Semi-supervised learning algorithms, on the other hand, use a combination of labeled (known) and unlabeled (unknown) data to improve their accuracy in identifying anomalies.

Pattern recognition in the context of IDS involves analyzing network traffic data to identify signatures or patterns indicative of malicious activities. AI models, especially deep learning models like Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN), excel at identifying patterns in complex and large datasets. These models can be trained on historical data containing various types of cyber-attacks and intrusion attempts [16]. As the AI learns the signatures of known attack vectors, it can recognize similar patterns in real-time network traffic data, alerting security teams to potential threats. Furthermore, AI models can also generalize patterns learned from historical data to identify new, previously unseen attack vectors that share similarities with known threats.

The most impactful role that these AI models can have is reducing false positives through advanced data analysis and correlation techniques. Moreover, AI models can be retrained and fine-tuned over time using feedback from security analysts, continually improving their ability to identify genuine threats and further reducing the number of false positives [17].

6. Case Studies and Empirical Evidence

In this analysis, we explored successful AI-powered intrusion detection system (IDS) applications across various sectors, emphasizing using artificial intelligence methods, such as neural networks and machine learning, to boost the identification and deterrence of cyber attacks.

6.1 Banking and Financial Services

A remarkable instance of efficient AI-driven IDS deployment is observed in banking and financial services. Kanimozhi and Dr. T. Prem Jacob conducted a study

[18] suggesting an Artificial Neural Network-oriented system for identifying botnet attacks, a significant risk to these sectors. The system was trained using the CSE-CIC-IDS2018 dataset, supplied by the Canadian Institute for Cybersecurity, and implemented on Amazon Web Services (AWS). The outcomes showed an unprecedented accuracy rate of 99.97%, an average ROC curve area of 0.999, and a minimal false positive rate of 0.03%. The success of this system emphasizes AI-driven IDS's potential to scrutinize network traffic and spot cyber threats in real time.

6.2 AI-Enhanced Honeypots

Scientists at the University of Texas at Dallas created DeepDig, an AI-integrated honeypot system that learns from assaults and morphs genuine network resources into lures [19]. This method addresses static deception technology limitations, which do not learn from previous attacks, leaving them vulnerable to AI-capable opponents. DeepDig employs machine learning approaches to gain a deeper insight into attackers' actions, improving the system's adaptability and defense against evolving hazards. By incorporating real assets into the honeypot, even the most skilled adversaries cannot escape interaction with the trap, enabling the IDS to learn and strengthen its defenses over time

6.3 Deep Learning in Network Intrusion Detection

Xu et al.'s case study, cspecc.utsa.edu, examined deep learning techniques' application for supervised network intrusion detection and unsupervised network anomaly detection [20]. The research methodically assessed deep learning's effectiveness in network intrusion detection, showcasing AI-powered techniques' potential in analyzing and pinpointing malicious traffic in real-time. This investigation adds to the expanding knowledge base on AI-driven IDS and encourages the creation of more advanced systems for protection against cyber threats.

7. Future Directions

Steps were taken in the initial direction of this paper to implement a ChatGPT-based intrusion detection system. This was done using OpenAI’s GPT-4 API [21]. An explanation of the theorized design, theory, and limitations will be addressed.

7.1 ChatGPT IDS Design

In making an AI IDS, GPT-4 was decided to be used as the model. This is due to its broad knowledge of domains, accurate problem-solving skills, and ability to complete complicated instructions. An existing network traffic capture tool or IDS can be used: in this research case, it was decided to use tcpdump. Tcpdump is a low-profile command-line packet analyzer that can export network traffic in pcap and CSV format [22]. The next step is to integrate GPT-4 into the existing network traffic capture tool or run it independently to analyze network traffic. GPT-4 would then analyze the incoming packet data and scan it for malicious activity. Because this language model is connected to the internet, it can scrape current websites for current threats and payloads for comparison with the incoming network traffic. This gives the AI intrusion detection program the added benefit of detecting newer threats that need to be constantly patched in existing IDS. The flow and analysis of data is shown in Figure 1.

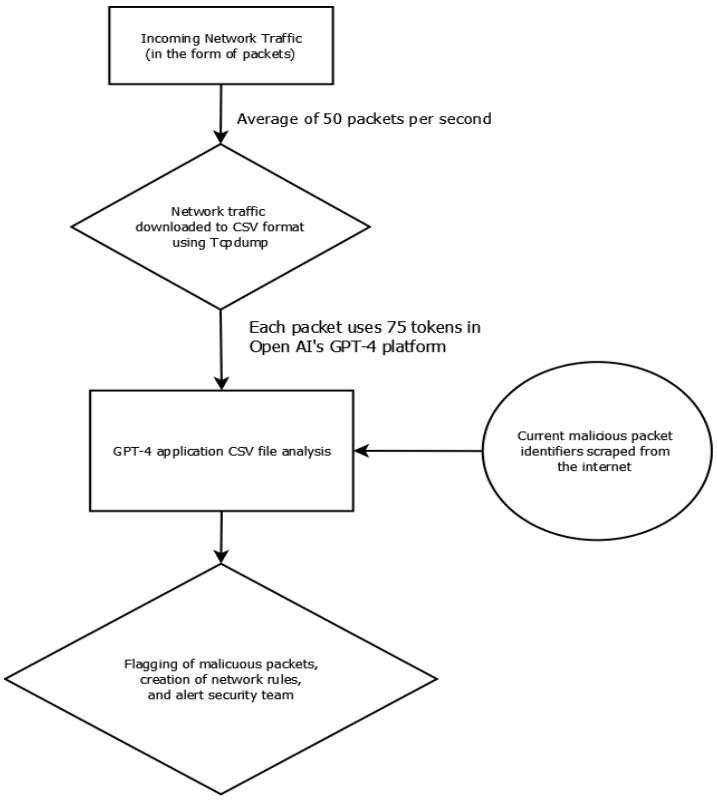

Figure 1: Data flow of proposed GPT-4 based IDS
Source: Michal Markevych

7.2 Design Limitations

The design in Figure 1 was constructed using a small model which sampled small amounts of data (under 50 packets). In order to analyze more packets, a large-language model must be used, to convert the data from the packets into a vector database. Vector databases store data and export it to be analyzed by programs such as AI models. This constrained the construction of this design as vector databases prices scale with the amount of data analyzed.

Another challenge came with GPT-4 API pricing, which is priced at .03 cents per 1000 tokens. Let us consider a regular network receiving 50 packets per second, and each packet requires 75 tokens to analyze (found using data to token calculator). Implementing this design costs $6.75 per minute of running the IDS system. This constraint could be mitigated by receiving access to OpenAI's GPT-4 API waitlist, which allows users to get more access to GPT-4 queries.

8. Conclusions

ChatGPT or similar AI models offer immense potential in significantly upgrading IDS. By incorporating AI algorithms, these enhanced systems provide heightened adaptability, superior pattern recognition, and precise real-time detection and response capabilities. This allows IDS to stay ahead of the constantly evolving threat landscape and deliver a more robust, proactive defense against cyber threats. However, despite these benefits, challenges, and areas still require further research. Enhancing the accuracy of AI-powered IDS and minimizing false positives remain crucial concerns.

As the field of AI and machine learning continues to advance, IDS is anticipated to become increasingly effective and efficient, further strengthening their ability to safeguard networks and computer systems from a broad spectrum of cyberthreats. Rising cybersecurity threats require modern solutions to protect critical infrastructure; this is a model to accomplish this task [23].

References List

[1] Delamore B., Ko R.K.L. Chapter 9 - Security as a service (SecaaS)—An overview [Internet]. Ko R, Choo KKR, editors. ScienceDirect. Boston: Syngress; 2015 [cited 2023 May 15]. p. 187–203. Available from: https://www.sciencedirect.com/science/article/abs/pii/B9780128015957000094

[2] Niksefat S., Kaghazgaran P., Sadeghiyan B. Privacy issues in intrusion detection systems: A taxonomy, survey and future directions. Computer Science Review. 2017 Aug;25:69–78.

[3] Aljanabi M., Ismail M.A., Ali A.H. Intrusion Detection Systems, Issues, Challenges, and Needs. International Journal of Computational Intelligence Systems. 2021;

[4] Aljanabi M., Ismail M.A., Ali A.H. Intrusion Detection Systems, Issues, Challenges, and Needs. International Journal of Computational Intelligence Systems. 2021;

[5] Liao H.J., Richard Lin C.H., Lin Y.C., Tung K.Y. Intrusion detection system: A comprehensive review. Journal of Network and Computer Applications [Internet]. 2013 Jan;36(1):16–24. Available from: https://www.sciencedirect.com/science/article/pii/S1084804512001944

[6] Cybersecurity Spotlight – Signature-Based vs Anomaly-Based Detection [Internet]. CIS. Available from: https://www.cisecurity.org/insights/spotlight/cybersecurity-spotlight-signature-based-vs-anomaly-based-detection

[7] Repalle S, Ratnam Kolluru V. Intrusion Detection System using AI and Machine Learning Algorithm. International Research Journal of Engineering and Technology.

[8] Li W., Yi P., Wu Y., Pan L., Li J. A New Intrusion Detection System Based on KNN Classification Algorithm in Wireless Sensor Network. Journal of Electrical and Computer Engineering [Internet]. 2014 [cited 2019 Nov 24];2014:1–8. Available from: https://www.hindawi.com/journals/jece/2014/240217/

[9] Sommer R., Paxson V. Outside the Closed World: On Using Machine Learning for Network Intrusion Detection. 2010 IEEE Symposium on Security and Privacy [Internet]. 2010 [cited 2019 Dec 6]; Available from: https://ieeexplore.ieee.org/abstract/document/5504793/

[10] Nobakht M., Sivaraman V., Boreli R. A Host-Based Intrusion Detection and Mitigation Framework for Smart Home IoT Using OpenFlow. 2016 11th International Conference on Availability, Reliability and Security (ARES). 2016 Aug;

[11] Jagadish H.V., Gehrke J., Labrinidis A., Papakonstantinou Y., Patel J.M., Ramakrishnan R., et al. Big data and its technical challenges. Communications of the ACM. 2014 Jul 1;57(7):86–94.

[12] Valdovinos I., Perez-Diaz J., Choo K.K., Botero J. Emerging DDoS attack detection and mitigation strategies in software-defined networks: Taxonomy, challenges and future directions. Journal of Network and Computer Applications [Internet]. 2021 Aug 1 [cited 2021 Sep 23];187:103093. Available from: https://www.sciencedirect.com/science/article/pii/S1084804521001156

[13] Drewek-Ossowicka A., Pietrołaj M., Rumiński J. A survey of neural networks usage for intrusion detection systems. Journal of Ambient Intelligence and Humanized Computing. 2020 May 12;12(1):497–514.

[14] Laghrissi F., Douzi S., Douzi K., Hssina B. IDS-attention: an efficient algorithm for intrusion detection systems using attention mechanism. Journal of Big Data. 2021 Nov 29;8(1).

[15] Khraisat A., Gondal I., Vamplew P., Kamruzzaman J. Survey of intrusion detection systems: techniques, datasets and challenges. Cybersecurity [Internet]. 2019 Jul 17;2(1). Available from: https://cybersecurity.springeropen.com/articles/10.1186/s42400-019-0038-7

[16] Otoum Y., Nayak A. AS-IDS: Anomaly and Signature Based IDS for the Internet of Things. Journal of Network and Systems Management. 2021 Mar 4;29(3).
[17] Kim A., Park M., Lee D.H. AI-IDS: Application of Deep Learning to Real-Time Web Intrusion Detection. IEEE Access. 2020;8:70245–61.
[18] Kanimozhi V., Jacob T.P. Artificial Intelligence based Network Intrusion Detection with hyper-parameter optimization tuning on the realistic cyber dataset CSE-CIC-IDS2018 using cloud computing. ICT Express. 2019 Apr;
[19] William D. How AI can help improve intrusion detection systems [Internet]. GCN. Available from: https://gcn.com/cybersecurity/2020/04/how-ai-can-help-improve-intrusion-detection-systems/291266/
[20] Fernández G., Xu S. A Case Study on Using Deep Learning for Network Intrusion Detection [Internet]. [cited 2023 May 15]. Available from: https://cspecc.utsa.edu/publications/files/Xu_2019_Case_Study_Deep_Learning_Net_Intr_Detect.pdf
[21] OpenAI. OpenAI [Internet]. OpenAI. 2019. Available from: https://openai.com/
[22] tcpdump. TCPDUMP/LIBPCAP public repository. Tcpdumporg [Internet]. 2017; Available from: https://www.tcpdump.org
[23] Dawson M., Bacius R., Gouveia L.B., & Vassilakos A. (2021). Understanding the challenge of cybersecurity in critical infrastructure sectors. Land Forces Academy Review, 26(1), 69-75.