International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347 - 6435 (Online) Volume 13, Issue 6, June 2024)

Review of Intrusion Detection Model for Prediction of Cyberattacks Using Artificial Intelligence Technique

[1]Aman Kumar Yadav, [2]Dr P. K. Sharma
[1]M.Tech Scholar, Department of Computer Science Engineering, NRI Institute of Research & Technology, Bhopal, India
[2]Principal, Department of Computer Science Engineering, NRI Institute of Research & Technology, Bhopal, India

***Abstract*— The rapid advancement of digital technologies has significantly increased the complexity and scale of cyber threats, necessitating robust and adaptive security mechanisms. This paper reviews intrusion detection models that leverage artificial intelligence (AI) techniques for predicting cyberattacks. It examines various AI methodologies, including machine learning (ML), deep learning (DL), and hybrid approaches, analyzing their effectiveness in identifying and mitigating sophisticated cyber threats. The review highlights key challenges in current intrusion detection systems (IDS) and proposes potential enhancements to improve detection accuracy and response times. Emphasis is placed on the integration of AI techniques with traditional IDS frameworks, focusing on how these innovative models can anticipate and thwart emerging cyber threats in real-time. The paper aims to provide a comprehensive understanding of the state-of-the-art AI-driven IDS, their current limitations, and future research directions to bolster cybersecurity defenses.**

***Keywords*—AI, IOT, Cyber, NIDS, HIDS, Security.**

I. Introduction

In the contemporary digital landscape, the proliferation of interconnected devices and systems has led to an unprecedented surge in cyber threats. Cyberattacks have evolved in complexity and frequency, targeting critical infrastructure, financial institutions, and individual users alike. Traditional security mechanisms, while foundational, are increasingly inadequate to counteract these sophisticated threats. This inadequacy underscores the need for more advanced and adaptive security solutions, specifically Intrusion Detection Systems (IDS) that can predict and mitigate cyber threats in real-time.

Intrusion Detection Systems play a pivotal role in the defense against cyberattacks by monitoring network traffic, identifying potential threats, and alerting administrators to suspicious activities. However, the static and signature-based nature of traditional IDS often fails to detect novel and advanced persistent threats (APTs). This limitation has driven the integration of Artificial Intelligence (AI) techniques into IDS, transforming them into more dynamic and intelligent systems capable of learning from data and identifying patterns indicative of cyber threats.

AI, particularly Machine Learning (ML) and Deep Learning (DL), has shown remarkable potential in enhancing IDS capabilities. ML algorithms can analyze vast amounts of data to detect anomalies and recognize patterns that signify potential intrusions. DL, a subset of ML, leverages neural networks with multiple layers to model complex data representations, improving the accuracy of threat detection. Hybrid approaches that combine multiple AI techniques have also been explored to leverage the strengths of different models and provide more comprehensive security solutions.

This review aims to provide an in-depth examination of AI-driven intrusion detection models for cyberattack prediction. It will explore various AI methodologies applied to IDS, including supervised learning, unsupervised learning, reinforcement learning, and hybrid models. The review will assess the effectiveness of these approaches in different scenarios, considering factors such as detection accuracy, false positive rates, scalability, and response times.

Furthermore, the paper will discuss the challenges and limitations associated with implementing AI in IDS, such as the need for large labeled datasets, the risk of adversarial attacks, and the computational complexity of AI models. It will also highlight recent advancements in the field, including the use of generative adversarial networks (GANs) for IDS training, the role of explainable AI in enhancing the interpretability of detection models, and the potential of federated learning for improving data privacy in IDS.

By providing a comprehensive overview of the current state of AI-driven IDS, this paper aims to contribute to the ongoing research and development in cybersecurity. It will identify key areas for future research, suggesting how AI can be further leveraged to develop more robust, adaptive, and proactive intrusion detection systems. The ultimate goal is to enhance the ability of IDS to anticipate and mitigate cyber threats, thereby strengthening the overall security posture of digital infrastructures.

II. LITERATURE SURVEY

The Intrusion Detection System (IDS) concept that was introduced by S. Ho et al., [1] is designed to identify network intrusions by categorizing all of the packet traffic in the network as either benign or malicious. Training and validation of the suggested model were carried out with the use of the dataset that was provided by the Canadian Institute for Cybersecurity Intrusion Detection System (CICIDS2017). A number of metrics, including the overall accuracy, attack detection rate, false alarm rate, and training overhead, have been analyzed and assessed for the model. A comparative analysis of the performance of the suggested model in comparison to nine other well-known classifiers has been published.

It is possible to identify such incursions with the use of datasets and by continuously updating them, as stated by V. K. Navya et al. [2]. One technique that stands out is the Deep Neural Network (DNN), which is a sort of deep learning model. This algorithm contributes to the development of an Intrusion Detection System (IDS) that is both flexible and effective, allowing it to identify and categorize intrusions that are unexpected and unanticipated.

In order to enhance the detection of cyberattacks, Y. A. Farrukh et al., [3] offer a two-layer hierarchical machine learning model that has an accuracy of 95.44%. The initial layer of the model is used to differentiate between the two modes of operation, which are classified as either the normal state or the assault. The second layer is that which is used for the purpose of categorizing the condition into various kinds of cyberattacks. With the layered method, the model is given the option to concentrate its training on the specific job that is being performed by the layer, which ultimately leads to an increase in the accuracy of the model. We compared the performance of the suggested model to that of other recent cyber attack detection models that have been presented in the literature in order to verify the efficacy of the model that was proposed.

According to S. Thirimanne et al.,, [4] the primary objective of this research is to identify the most effective machine learning algorithm for intrusion detection. This algorithm will be trained using the NSL-KDD and the UNSW-NB15 datasets. Additionally, a comparative analysis will be carried out between six different machine learning algorithms that are classified as supervised, semi-supervised, and unsupervised learning. Support Vector Machines (SVM) and Deep Neural Networks (DNN) perform better for NSL-KDD and UNSW-NB15, respectively, according to the findings of this study, which showed that the performance of supervised and semi-supervised machine learning algorithms outperformed the performance of unsupervised machine learning algorithms for both datasets.

An overview of DRL strategies that have been developed for cyber security is presented by T. T. Nguyen et al., [5]. We concentrate on many critical topics, including DRL-based security approaches for cyber-physical systems, autonomous intrusion detection techniques, and multiagent DRL-based game theory simulations for defensive tactics against cyberattacks. Extensive discussions and future research ideas on DRL-based cyber security are also included in this article. We anticipate that this exhaustive evaluation will provide the groundwork for future research on examining the potential of developing DRL to deal with more complex cyber security issues and will also make it easier for such research to be conducted.

The suggested model by W. Xu et al. [6] makes use of the most efficient reconstruction error function, which is a crucial component in the model's ability to determine whether a network traffic sample is typical or abnormal. Because of these sets of unique methodologies and the ideal model architecture, our model is able to be better prepared for feature learning and dimension reduction, which ultimately results in improved detection accuracy as well as f1-score. We assessed our suggested model on the NSL-KDD dataset, which outperformed other approaches that were comparable by attaining the greatest accuracy and f1-score in detection, which were respectively 90.61% and 92.26%.

The attention mechanism developed by K. Cao et al. [7] is applied in order to accurately collect essential qualities that are representative of the structural properties of traffic data. In addition, a CuDNN-based long short-term memory network is used in order to swiftly accelerate the convergence of the model while simultaneously learning time-related information on the traffic. Finally, global maxpooling is implemented in order to increase the generalization capabilities of the proposed model and to reduce the amount of data that is included inside it. The results of the experiments conducted on the UNSW-NB15 dataset demonstrate that the suggested model has an accuracy of binary classification that may reach up to 92.65%. In addition to that, it has an accuracy of 81.28% when it comes to identifying different types of assaults.

I. Ullah et al., [8], a multiclass classification model is developed with the help of a convolutional neural network model. After that, the suggested model is put into action by using convolutional neural networks in one-dimensional, two-dimensional, and three-dimensional space. The BoT-IoT, IoT Network Intrusion, MQTT-IoT-IDS2020, and IoT-23 intrusion detection datasets are used in order to evaluate the convolutional neural network model that has been presented. Through the use of a convolutional neural network multiclass pre-trained model, transfer learning is utilized to achieve binary and multiclass classification. The binary and multiclass classification models that we have developed have obtained great levels of accuracy, precision, recall, and F1 score when compared to the deep learning implementations that are now in development.

According to D. Park et al. [9], a model for an intrusion detection system that is based on deep learning has been developed. This model examines sophisticated attack patterns by performing data learning. Deep learning models, on the other hand, have the drawback of needing to teach themselves new information every time a new cyberattack strategy is discovered. The amount of time that is necessary to learn a substantial quantity of facts is not beneficial. Using the Leipzig Intrusion Detection Data Set (LID-DS), which is a host-based intrusion detection data set that was published in 2018, an experiment was carried out for the purpose of this study.

An outlier detection (also known as anomaly detection) problem and a challenging multiclass classification problem consisting of 14 classes (13 Modbus/TCP cyberattacks and normal instances) were successfully solved by the proposed intrusion detection system (IDS) that was developed by I. Siniosoglou et al., [10]. This IDS was validated in four real-world SG evaluation environments, which are as follows: (a) SG lab, (b) substation, (c) hydropower plant, and (d) power plant. Also, MENSA is able to differentiate between five different cyberattacks directed on DNP3.

III. CHALLENGES

Implementing artificial intelligence (AI) techniques in Intrusion Detection Systems (IDS) presents several challenges that need to be addressed to enhance their effectiveness in predicting and mitigating cyberattacks. These challenges encompass technical, operational, and ethical aspects, which are crucial for the successful deployment and operation of AI-driven IDS.

1. Data Quality and Availability

- **Large Labeled Datasets:** AI models, particularly those based on machine learning (ML) and deep learning (DL), require extensive labeled datasets to train effectively. However, acquiring such datasets can be challenging due to the scarcity of real-world attack data and the time-consuming nature of labeling data accurately.
- **Imbalanced Data:** In cybersecurity, attack data is often significantly less prevalent than normal data, leading to imbalanced datasets. This imbalance can cause AI models to be biased towards detecting non-malicious activities, resulting in higher false negative rates where actual threats go undetected.

2. Adversarial Attacks

- **Evasion Techniques:** Attackers continually develop new methods to evade detection by AI-driven IDS. Adversarial attacks, where attackers manipulate inputs to deceive AI models, pose a significant threat. These evasion techniques can undermine the reliability of AI models, making them less effective in real-world scenarios.
- **Model Robustness:** Ensuring the robustness of AI models against adversarial attacks is a major challenge. Models must be designed to withstand various types of attacks and maintain their detection accuracy under adversarial conditions.

3. Computational Complexity

- **Resource Intensive:** AI models, especially DL models, require substantial computational resources for training and inference. This includes powerful hardware such as GPUs and large memory capacities, which may not be feasible for all organizations.

- **Real-time Processing:** Implementing AI-driven IDS for real-time threat detection necessitates low-latency processing capabilities. Balancing the computational demands of AI models with the need for swift threat detection is a critical challenge.

4. Interpretability and Explainability

- **Black-box Nature:** Many AI models, particularly DL models, operate as black boxes, providing little insight into how they arrive at their decisions. This lack of transparency can hinder trust and adoption, as security analysts need to understand the reasoning behind detected threats.
- **Explainable AI:** Developing explainable AI models that provide clear, understandable explanations for their decisions is essential for gaining trust and facilitating effective responses to detected threats.

5. Integration with Existing Systems

- **Legacy Systems Compatibility:** Integrating AI-driven IDS with existing legacy security systems can be complex. Ensuring compatibility and seamless operation with traditional IDS and other security tools requires careful planning and execution.
- **Operational Disruptions:** Transitioning to AI-driven IDS may cause operational disruptions, necessitating thorough testing and gradual implementation to minimize impact on organizational workflows.

IV. CONCLUSION

The integration of artificial intelligence (AI) techniques into Intrusion Detection Systems (IDS) marks a significant advancement in the field of cybersecurity. This review highlights the transformative potential of AI-driven IDS in predicting and mitigating cyberattacks, emphasizing their superior capabilities in detecting sophisticated and evolving threats compared to traditional IDS. AI methodologies, including machine learning (ML), deep learning (DL), and hybrid approaches, have demonstrated enhanced accuracy and efficiency in identifying malicious activities and adapting to new attack vectors. Future research should focus on developing more robust and explainable AI models.

REFERENCES

1. S. Ho, S. A. Jufout, K. Dajani and M. Mozumdar, "A Novel Intrusion Detection Model for Detecting Known and Innovative Cyberattacks Using Convolutional Neural Network," in IEEE Open Journal of the Computer Society, vol. 2, pp. 14-25, 2021, doi: 10.1109/OJCS.2021.3050917.
2. V. K. Navya, J. Adithi, D. Rudrawal, H. Tailor and N. James, "Intrusion Detection System using Deep Neural Networks (DNN)," 2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA), 2021, pp. 1-6, doi: 10.1109/ICAECA52838.2021.9675513.
3. Y. A. Farrukh, Z. Ahmad, I. Khan and R. M. Elavarasan, "A Sequential Supervised Machine Learning Approach for Cyber Attack Detection in a Smart Grid System," 2021 North American Power Symposium (NAPS), 2021, pp. 1-6, doi: 10.1109/NAPS52732.2021.9654767.
4. S. Thirimanne, L. Jayawardana, P. Liyanaarachchi and L. Yasakethu, "Comparative Algorithm Analysis for Machine Learning Based Intrusion Detection System," 2021 10th International Conference on Information and Automation for Sustainability (ICIAfS), 2021, pp. 191-196, doi: 10.1109/ICIAfS52090.2021.9605814.
5. T. T. Nguyen and V. J. Reddi, "Deep Reinforcement Learning for Cyber Security," in IEEE Transactions on Neural Networks and Learning Systems, doi: 10.1109/TNNLS.2021.3121870.
6. W. Xu, J. Jang-Jaccard, A. Singh, Y. Wei and F. Sabrina, "Improving Performance of Autoencoder-Based Network Anomaly Detection on NSL-KDD Dataset," in IEEE Access, vol. 9, pp. 140136-140146, 2021, doi: 10.1109/ACCESS.2021.3116612.
7. K. Cao, J. Zhu, W. Feng, C. Ma, M. Liu and T. Du, "Network Intrusion Detection based on Dense Dilated Convolutions and Attention Mechanism," 2021 International Wireless Communications and Mobile Computing (IWCMC), 2021, pp. 463-468, doi: 10.1109/IWCMC51323.2021.9498652.
8. I. Ullah and Q. H. Mahmoud, "Design and Development of a Deep Learning-Based Model for Anomaly Detection in IoT Networks," in IEEE Access, vol. 9, pp. 103906-103926, 2021, doi: 10.1109/ACCESS.2021.3094024.
9. D. Park, S. Kim, H. Kwon, D. Shin and D. Shin, "Host-Based Intrusion Detection Model Using

Siamese Network," in IEEE Access, vol. 9, pp. 76614-76623, 2021, doi: 10.1109/ACCESS.2021.3082160.

10. I. Siniosoglou, P. Radoglou-Grammatikis, G. Efstathopoulos, P. Fouliras and P. Sarigiannidis, "A Unified Deep Learning Anomaly Detection and Classification Approach for Smart Grid Environments," in IEEE Transactions on Network and Service Management, vol. 18, no. 2, pp. 1137-1151, June 2021, doi: 10.1109/TNSM.2021.3078381.