JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Automated Cyber Attack Response Using Artificial Intelligence

Mrs. Dnyaneshwari Anil Bhor,
Mrs. Bhakti Chandrakant Kale (Assistant Professor)
Department of Master of Computer Application,
Second Year MCA
JCEI's Jaihind Institute of Management And Research Vadgaon Sahani, India(JIMR)
PUNE, MAHARASHTRA, INDIA

Abstract: The rapid development of internet networks, cloud service providers, and connected devices increases day by day the number and complexity of cyber-attacks. Conventional security systems are based primarily on manually defined rules and signatures, which may not always be sufficient for detecting and responding to emerging threats. Automated Cyber Attack Response Using Artificial Intelligence suggests a novel intelligent security framework that utilizes Artificial Intelligence (AI) and Machine Learning (ML) to provide automatic responses to cyber-attack incidents. Anomalies related to attack events are detected throughout the continuous monitoring of network traffic flows and system messages. Various attack categories, including denial-of-service (DoS) and brute-force attacks, malware activity, and unauthorized access attempts, can all be identified using machine learning algorithms. The proposed framework can automatically generate a security event message and take appropriate action in response to the detected attack, such as isolation of compromised devices or blocking malicious traffic. The given concept aims to improve the speed of incident response and reduce the workload of human analysts in the cyber security domain. Artificial intelligence helps cybersecurity systems learn from past experiences to improve their detection and response capabilities against different kinds of attacks. Automated Cyber Attack Response Using Artificial Intelligence can be evaluated using standard cybersecurity benchmark datasets and performance metrics such as accuracy, precision, recall, F1 – score, and detection time. The suggested approach illustrates how artificial intelligence can be applied to automate network traffic analysis and improve the cyber-attack detection and response capabilities of contemporary systems.

***Keyword*- Artificial Intelligence, Cyber Security, Machine Learning, Attack Detection, Automated Response, Network Security.**

I. INTRODUCTION

The advanced development of information technologies, cloud computing, Internet of Things(IoT), and digital communication has made computer networks an essential part of modern business. Various enterprises, educational facilities, medical organizations, financial institutions, government agencies, and other entities increasingly rely on computers and related technologies in order to store, process, and transmit significant amounts of data. However, the growing importance and use of interconnected systems and databases generate serious security concerns. The main issue is connected to the fact that attackers try to enter corporate networks illegally and unauthorized in order to steal data, disrupt operations, and damage computer hardware and software.

It is rather common for conventional cybersecurity systems and software to use rule-based and signature-based approaches for the detection of malicious activity, following specific pre-set instructions in order to mitigate and respond to attacks. While such methods have proven to be effective in terms of recognized threats, they might be inappropriate and ineffective in case of new and unknown threats. Moreover, time is often needed for administrators to analyze suspicious activity and take appropriate action, which in turn can lead to adverse consequences. A timely response to an attack is essential in order to minimize losses and operational damage.

AI can play a significant role in terms of improving cybersecurity. The main advantage of using AI is the ability to process and analyze vast amounts of network data in search of possible attacks and threats automatically. Specifically, Machine Learning can help cybersecurity systems and software detect various types of attacks, including but not limited to Denial-of-Service, DDoS, brute-force, worms, viruses, and zero-day. The key strength of using AI in cybersecurity is the possibility of automated processing and analysis of network traffic data in order to identify irregularities and suspicious patterns.

However, it is essential to recognize that detecting an attack is not sufficient, and cyber defense systems need to respond to it in real-time in order to eliminate the adverse impact. Automated Cyber Attack Response Using Artificial Intelligence aims at introducing

an intelligent system capable of analyzing the characteristics of an attack, determining its type, and taking appropriate action in order to neutralize and mitigate its impact. Particular attention is paid to the process of analyzing network traffic data and automating the response to cyber-attacks, which includes blocking and isolating unauthorized remote access, filtering network packets, and raising alerts.

It is possible to integrate AI into the cybersecurity domain in order to reduce human intervention and ensure rapid response to cyber incidents. It is possible to conduct a series of experiments in order to evaluate the effectiveness of the introduced methods, which will include the use of different types of security datasets and specific performance indicators. Thus, it will be possible to improve the overall state of cybersecurity and make it more effective and efficient in terms of responding to cyber incidents.

II. Problem Statement

Nowadays, due to the rapid rise in interconnected devices, cloud technologies, and the Internet, the number of cyber-attacks has grown significantly. The current cybersecurity solutions are mostly based on detecting known attacks through rule-based mechanisms, while the constantly changing nature of cyber threats makes it hard to come up with a solution that can identify and mitigate unknown attacks automatically. Cybersecurity analysts spend too much time investigating detected incidents, which leaves attackers enough time to inflict more damage. Hence, a system is required that is able to detect and respond to attacks promptly in order to minimize damage, time spent on investigation, and human interaction. In this research, a framework is proposed that is based on Artificial Intelligence (AI) and Machine Learning (ML) and is capable of analyzing network traffic data to identify suspicious behaviour, classify attacks, and take relevant mitigation actions automatically.

III. Objectives

The intent of "Automated Cyber Attack Response Using Artificial Intelligence" is to design a framework that detects and responds to cyber-attack automatically. The objectives include:

1. Developing an AI-based framework for detecting suspicious and malicious traffic activities.
2. Designing a framework that uses Machine Learning to classify and identify different cyber-attack types.
3. Analyzing the network traffic and events to detect abnormal behaviours.
4. Designing a framework that offers automated response mechanisms in detecting different attack activities.
5. Reducing the time taken when detecting and responding to attacks and threats.

IV. Literature Review

The following literature review is appropriate for the research topic "Automated Cyber Attack Response Using Artificial Intelligence." The articles discussed below are related to intrusion detection, machine learning, deep learning, anomaly detection, and automated/AI-assisted cybersecurity. Based on the findings, it could be stated that AI could be used for improving attack detection capabilities. However, there are several issues associated with using such technologies in practice, including response time, false-positive identification, evolving patterns of attacks, training data quality, explainability, and operational safety.

No.	Author(s) & Year	Existing Work / Method	Main Limitation
1	Dorothy E. Denning (1987)	suggested a model for intrusion detection that relies on spotting unusual activity in computer systems.	The strategy was primarily conceptual and ignored automatic response or contemporary high-volume network traffic.
2	W. Lee & S. J. Stolfo (2000)	created a framework for employing data-mining techniques to create features and models for intrusion detection.	The quality of training data has a significant impact on feature engineering and model performance.
3	R. Lippmann et al. (2000)	used several detection techniques and DARPA evaluation datasets to assess network intrusion detection.	The complex and quickly evolving network environments of today are not adequately represented by the evaluation datasets.
4	Mukkamala, Janoski & Sung (2005)	investigated intrusion detection methods using neural networks and support vector machines.	Practical deployment may be impacted by computational constraints and reliance on particular features.
5	Moustafa & Slay (2015)	presented the UNSW-NB15 dataset, which includes both harmful and typical network traffic from today.	It is still a benchmark dataset and might not accurately reflect all real-world network environments, while being more modern than prior datasets. Predefined training and testing sets are provided by the dataset. (UNSW Websites)
6	Buczak & Guven (2016)	provided a thorough analysis of machine learning and data mining techniques for cybersecurity intrusion detection.	The study examined detection rather than full automatic response and found issues with data quality, algorithm complexity, and the use of appropriate machine learning techniques. (Xplore by IEEE)

No.	Author(s) & Year	Existing Work / Method	Main Limitation
7	Javaid et al. (2016)	studied the use of neural networks for attack classification and applied deep learning techniques to network intrusion detection.	Appropriate training data and computer power are necessary for deep learning, yet generalization to new settings is still difficult.
8	Kim et al. (2016)	suggested an LSTM recurrent neural network classifier for intrusion detection that finds attack patterns using sequential learning.	Training and tuning recurrent models can be computationally demanding, and performance depends on representative sequential data.
9	Tang et al. (2016)	suggested a deep learning method for Software Defined Networking (SDN) network intrusion detection.	Since the approach was designed for an SDN context, its findings might not apply to other network topologies.
10	Yin et al. (2017)	suggested a deep learning intrusion detection method that classifies attacks using Recurrent Neural Networks (RNNs). .	Without appropriate retraining, RNN models may struggle to adjust to new attack types and may require substantial training resources.
11	Hodo et al. (2017)	examined and contrasted the relevance of deep learning and neural-network techniques for intrusion detection to cybersecurity.	Attack categories, feature selection, and dataset features can all affect detection performance.
12	Mirsky et al. (2018)	Kitsune is a proposed online network intrusion detection system that employs a group of autoencoders. It was proved to function even on hardware with limited resources and was intended for unsupervised, online detection. (arXiv)	The system's main goal is to identify anomalous behavior; choosing and carrying out the proper protective measures on its own is a different task.
13	Shone et al. (2018)	suggested a deep-learning network intrusion detection method that uses Random Forest for classification and non-symmetric deep autoencoders for feature extraction.	On earlier datasets, the approach demonstrated issues with minority attack classes like R2L and U2R. (MDPI)
14	Sharafaldin, Lashkari & Ghorbani (2018)	created the CICIDS2017 dataset for IDS research, which includes both benign traffic and a number of modern attack scenarios.	Experiments based on datasets might not be able to capture every variation of real-world network traffic and constantly changing threats.
15	Khairat et al. (2019)	provided a thorough analysis of IDS methods, datasets, attack kinds, and difficulties, including anomaly-based and signature-based detection. (Springer)	The study noted persistent issues such evasion strategies, zero-day attacks, dataset constraints, and precise detection.
16	Vinaykumar et al. (2019)	evaluated a number of deep-learning techniques for host and network-level intelligent intrusion detection. The study focused on automated attacks. detection and classification. (University of East London Repository)	The authors emphasized the difficulty of evolving assaults, the volume of data, and the requirement for scalable solutions. (University of East London Repository)
17	Alom et al. (2019)	examined neural-network-based deep learning methods for cybersecurity and intrusion detection. .	When compared to more straightforward machine learning models, deep learning models may use more resources and be less interpretable.
18	Ferrag et al. (2020)	examined cybersecurity datasets, deep learning techniques, and comparative intrusion detection techniques. The study looked at 35 cybersecurity datasets in a number of different categories. (ScienceDirect)	It is challenging to directly compare IDS approaches due to differences in datasets, metrics, preprocessing methods, and experimental settings.
19	Khan et al. (2021)	examined deep learning and machine learning methods for automatically identifying malicious traffic and detecting network intrusions.	When traffic characteristics deviate from the training data, the model's performance may suffer; false positives are still a major worry.
20	Dalmet (2026)	examined the use of machine learning, deep learning, reinforcement learning, XAI, SOAR, and related techniques for AI-driven threat identification and automated response. In particular, the study links automated incident response and AI-based detection. (Forensics for Digital Security)	Zero-day detection, adversarial assaults, explainability, safe automation, and the integration of AI with security operations are among the persistent problems noted in the review.

The examined research demonstrates a distinct shift from conventional intrusion detection to security systems based on machine learning and deep learning. Finding unusual or illegal activity was the primary focus of early studies. Subsequent research developed machine learning methods for automated attack classification, including SVM, Random Forest, Decision Tree, and Neural Networks. CNN, RNN, LSTM, autoencoders, and other deep-learning techniques have been used in more recent studies to enhance the detection of intricate network attacks. (DOI)

However, most previous studies concentrate on attack detection and categorization rather than complete automated response. Even if a detection system finds an attack, a security administrator must still determine the best course of action. As a result, there is a delay between identifying threats and taking prompt action.

V. Research Gap

The majority of current cybersecurity research has been on employing machine learning and deep learning approaches to identify and categorize cyberattacks. Nevertheless, after an assault is identified, a lot of systems still rely on human intervention. There is little research that integrates intelligent decision-making, automatic reaction, and real-time assault detection into a unified framework. Other issues that current methods may encounter include false alarms, shifting attack patterns, small datasets, and challenges in identifying unidentified attacks. An AI-based system that can recognize suspicious activity, determine the sort of attack, and automatically carry out appropriate response actions is therefore required. By creating an automated cyberattack response system using artificial intelligence that can speed up response times and enhance network security, the proposed research seeks to close this gap.

VI. Proposed Methodology

The "Automated Cyber Attack Response Using Artificial Intelligence" methodology is intended to identify cyberattacks, categorize threats, and automatically carry out appropriate countermeasures. The steps of the suggested framework are as follows:

1. Information Gathering

Data on network traffic and security events is gathered from appropriate cybersecurity datasets, such as UNSW-NB15, CICIDS2017, or other publically accessible datasets. Both typical network activity and other cyberattacks are included in the data.

2. Preprocessing of Data

Before the AI model is trained, the gathered data is cleansed. Noisy data, irrelevant features, duplicate records, and missing values are all addressed. When necessary, feature scaling is implemented once categorical values are transformed into numerical form.

3. Selection and Extraction of Features

Key network characteristics are chosen, including source and destination IP addresses, port numbers, protocols, packet sizes, connection durations, and traffic frequencies. Model performance is enhanced and superfluous information is reduced with the aid of feature selection.

4. AI-Powered Attack Identification

Network activity can be classified as malicious or normal using machine learning methods like Random Forest, Support Vector Machine, Decision Tree, or XGBoost. For intricate assault patterns, deep learning models may also be taken into consideration.

5. Classification of Attacks

The system categorizes the assault into groups like DoS/DDoS, brute-force, malware, scanning, or illegal access after identifying suspicious activities.

6. Automated Decision-Making

Based on the attack's severity and detection, the system chooses the best course of action. Blocking harmful communications, cutting off dubious connections, isolating impacted devices, or sending out alerts are all possible courses of action.

7. Automated Reaction Processing

Security features like firewall rules, access-control mechanisms, or network isolation automatically apply the chosen response. As a result, the interval between attack detection and mitigation is shortened.

8. Assessment of Performance

Accuracy, precision, recall, F1-score, false-positive rate, detection time, and response time are used to assess the suggested system.

Proposed Workflow

Data Collection → Preprocessing → Feature Selection → AI-Based Detection → Attack Classification → Threat Severity Analysis → Automated Response → Performance Evaluation

Conclusion

An approach to enhancing cybersecurity through intelligent attack detection and automated response is offered by the suggested Automated Cyber Attack Response Using Artificial Intelligence framework. The system analyzes network activity, recognizes anomalous behavior, and categorizes various cyberattacks using artificial intelligence and machine learning techniques. The framework can start appropriate reaction activities, including blocking malicious traffic, isolating impacted systems, or creating notifications, after identifying a threat. This method can lessen reliance on manual security procedures and speed up response times. However, it is necessary to take into account issues like false positives, shifting assault patterns, a lack of training data, and safe automated decision-making. Future research can concentrate on explainable AI, adaptive learning, real-time implementation, zero-day threat detection, and interaction with security orchestration systems to increase the efficacy.

References

.

[1] D. E. Denning, “An intrusion-detection model,” IEEE Transactions on Software Engineering, vol. SE-13, no. 2, pp. 222–232, 1987.

[2] W. Lee, S. J. Stolfo, and K. W. Mok, “A data mining framework for building intrusion detection models,” in Proc. IEEE Symp. Security and Privacy, 1999, pp. 120–132.

[3] A. L. Buczak and E. Guven, “A survey of data mining and machine learning methods for cyber security intrusion detection,” IEEE Communications Surveys & Tutorials, vol. 18, no. 2, pp. 1153–1176, 2016. (DOI)

[4] R. Sommer and V. Paxson, “Outside the closed world: On using machine learning for network intrusion detection,” in Proc. IEEE Symp. Security and Privacy, 2010, pp. 305–316.

[5] N. Moustafa and J. Slay, “UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set),” in Proc. Military Communications and Information Systems Conference (MilCIS), 2015, pp. 1–6, doi: 10.1109/MilCIS.2015.7348942. (Adelaide Digital Library)
[6] N. Moustafa and J. Slay, “The evaluation of network anomaly detection systems: Statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set,” Information Security Journal: A Global Perspective, vol. 25, no. 1–3, pp. 18–31, 2016, doi: 10.1080/19393555.2015.1125974. (Taylor & Francis Online)
[7] T. T. T. Nguyen and G. Armitage, “A survey of techniques for internet traffic classification using machine learning,” IEEE Communications Surveys & Tutorials, vol. 10, no. 4, pp. 56–76, 2008.
[8] J. Kim, J. Kim, H. L. T. Thu, and H. Kim, “Long short-term memory recurrent neural network classifier for intrusion detection,” in Proc. Int. Conf. Platform Technology and Service (PlatCon), 2016, pp. 1–5.
[9] T. T. Tang, M. Mhamdi, D. McLernon, S. A. R. Zaidi, and A. Ghogho, “Deep learning approach for network intrusion detection in software defined networking,” in Proc. IEEE Int. Conf. Wireless Networks and Mobile Communications, 2016, pp. 258–263.
[10] N. Shone, T. N. Ngoc, V. D. Phai, and Q. Shi, “A deep learning approach to network intrusion detection,” IEEE Transactions on Emerging Topics in Computational Intelligence, vol. 2, no. 1, pp. 41–50, 2018.
[11] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, “Toward generating a new intrusion detection dataset and intrusion traffic characterization,” in Proc. Int. Conf. Information Systems Security and Privacy (ICISSP), 2018, pp. 108–116.
[12] Y. Mirsky, T. Doitshman, Y. Elo vici, and A. Shabtai, “Kitsune: An ensemble of autoencoders for online network intrusion detection,” in Proc. Network and Distributed System Security Symposium (NDSS), 2018.
[13] A. Javaid, Q. Niyaz, W. Sun, and M. Alam, “A deep learning approach for network intrusion detection system,” in Proc. 9th EAI Int. Conf. Bio-inspired Information and Communications Technologies, 2016, pp. 21–26.
[14] C. Yin, Y. Zhu, J. Fei, and X. He, “A deep learning approach for intrusion detection using recurrent neural networks,” IEEE Access, vol. 5, pp. 21954–21961, 2017.
[15] R. Vinaykumar, M. Alazab, K. P. Soman, P. Poornachandran, A. Al-Nimrat, and S. Venkatraman, “Deep learning approach for intelligent intrusion detection system,” IEEE Access, vol. 7, pp. 41525–41550, 2019.
[16] A. Khraisat, I. Gondal, P. Vamplew, and J. Kamruzzaman, “Survey of intrusion detection systems: Techniques, datasets and challenges,” Cybersecurity, vol. 2, Art. no. 20, 2019.
[17] N. Moustafa, G. Creech, and J. Slay, “Novel geometric area analysis technique for anomaly detection using trapezoidal area estimation on large-scale networks,” IEEE Transactions on Big Data, vol. 5, no. 4, pp. 481–494, 2019. (ETASR)
[18] Y. Mirsky, B. Shabtai, Y. Elovici, L. Rokach, and A. Shabtai, “Intelligent intrusion detection systems: A survey,” Artificial Intelligence Review, 2023.
[19] M. A. Ferrag, L. Maglaras, S. Moschoyiannis, and H. Janicke, “Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study,” Journal of Information Security and Applications, vol. 50, Art. no. 102419, 2020.
[20] M. A. Habeeb, M. S. K. Khan, and others, “Network intrusion detection system: A survey on artificial intelligence-based techniques,” Expert Systems, 2022. (Wiley Online Library)