International Journal of Applied and Advanced Multidisciplinary Research (IJAAMR)
Vol. 2, No. 1, 2024: 13-22

Intrusion Detection Systems Using Machine Learning

Anuja Phapale[1]*, Rohit Utekar[2],
Department of Information Technology, AISSMS IOIT
Corresponding Author: Rohit Utekar rohitutekar412@gmail.com

A R T I C L E I N F O

Keywords: Machine Learning, Intrusion, Detection, Supervised, Unsupervised, Dataset

Received : 25, November
Revised : 20, December
Accepted: 22, January

A B S T R A C T

The utilization of machine learning to enhance Intrusion Detection Systems (IDS). It encompasses an exploration of diverse IDS categories, fundamental evaluation metrics, and the dynamic landscape of machine learning methodologies. Recent trends underscore a shift towards the adoption of deep learning techniques for improving attack detection capabilities. Challenges arise from heightened model complexity and increased resource requirements. The paper also suggests future directions that encompass the development of updated datasets and the efficient management of resources through cloud integration. Throughout, this study emphasizes the continuous demand for research and innovation in the field of cybersecurity.

INTRODUCTION

The introduction of an Intrusion Detection System (IDS) is a fundamental step in bolstering network security and safeguarding digital assets. An IDS is a critical component of an organization's cybersecurity infrastructure, designed to monitor, detect, and respond to unauthorized and malicious activities within a network or system. Its primary purpose is to identify and alert on potential security breaches, intrusions, or suspicious behaviour, helping security teams respond swiftly and effectively to mitigate threats. This proactive approach to cybersecurity is essential in an era marked by an ever-evolving threat landscape, where cyberattacks continue to grow in complexity and frequency. The introduction of an IDS serves as a vigilant guardian, enhancing an organization's resilience against cyber threats and contributing to a robust defence strategy in the digital age.

Intrusion Detection Systems (IDS)

Organizations may monitor and safeguard their computer networks and systems against malicious activity, illegal access, and security risks with the aid of intrusion detection systems (IDS), which are essential components of cyber security. IDS are essential for seeing possible security issues and taking immediate action or using post-event analysis to address them. IDS comes in two primary flavors: host-based IDS (HIDS) and network-based IDS (NIDS).

Network-based IDS (NIDS):

- NIDS are deployed at strategic points within a network, such as at network gateways or on network segments.
- They analyse network traffic in real-time, looking for suspicious patterns, signatures, or anomalies.
- NIDS can detect known attack patterns based on pre-defined signatures or behavioural anomalies.
- When suspicious activity is detected, NIDS can generate alerts and trigger responses, such as logging, blocking, or notifying administrators.
- Examples of NIDS include Snort, Suricata, and Bro/Zeek.

Host-based IDS (HIDS):

- HIDS are installed on individual hosts or servers and monitor the activities and configurations of those specific systems.
- They can detect suspicious changes in system files, registry settings, and user account activities.
- HIDS are particularly effective at detecting insider threats and attacks that originate from within the network.
- Similar to NIDS, HIDS can generate alerts and take actions when suspicious activities are identified.
- Examples of HIDS include OSSEC, Tripwire, and McAfee Host Intrusion Prevention System (HIPS).

Key features and functions of IDS include:

- Signature-Based Detection: IDS utilize predefined signatures or patterns of known attacks to recognize malicious activities.
- Anomaly-Based Detection: Some IDS leverage machine learning and behavioural analysis to identify deviations from typical network or system behaviour.
- Real-Time Monitoring: IDS continually oversee network traffic or host activities to promptly detect threats as they occur.
- Alerts and Notifications: When suspicious activity is detected, IDS generate alerts to notify administrators or security personnel.
- Logging and Reporting: IDS record information about identified incidents, supplying valuable data for post-incident analysis and compliance reporting.

LITERATURE REVIEW

Andresini, et all said proposed a solution that combines an unsupervised approach with two auto encoders and a supervised stage for dataset construction. They initially trained the two auto encoders separately using normal and attack traffic. Subsequently, these auto encoders reconstructed the samples, which were then incorporated into the dataset used for model training. This dataset underwent one-dimensional CNN processing to assess the impact of one channel on the other, with the aim of enhancing differentiation between the two classes: normal and attack. Finally, a Soft-max classifier was employed to determine whether the data represented an attack or normal activity. Their model was tested on three datasets: KDDcup99, UNSW-NB15, and CICIDS2017, resulting in respective overall accuracies of 92.49%, 93.40%, and 97.90%. However, a limitation of their solution is the lack of detailed information regarding the various types of attacks."

Ali, et all proposed a model that utilizes a Fast Learning Network (FLN) based on particle swarm optimization (PSO). They employed PSO to enhance the accuracy of FLN, as FLN can be inefficient due to the weights used in the neural network. Their solution was assessed using the KDDcup99 dataset in comparison to other FLN-based solutions. They achieved a higher accuracy in distinguishing between different classes compared to other solutions, resulting in an overall accuracy of 89.23%. However, their overall accuracy was negatively impacted by their lower accuracy in identifying one of the smaller attack classes (R2L)."

Dong, et al**.** proposed a hybrid solution that combines clustering with SVM. In their approach, they first applied K-means clustering to preprocess the data, dividing it into distinct subsets. They subsequently used SVM on each of these subsets. Their solution was tested using the NSL-KDD datasets, resulting in an impressive overall accuracy of 99.45%. Moreover, when compared to other methods, their solution exhibited enhanced detection rates and also required less processing time compared to SVM algorithms that use different

parameters. However, it is important to note that the authors did not provide specific information regarding the accuracy of each attack classification."

Wisanwanichthan, et all proposed a Double- Layered Hybrid Approach (DLHA). In their solution, they began by segmenting the NSL-KDD dataset into two distinct groups. The first group encompassed all classes, while the second group included only the U2R, R2L, and normal classes. This division aimed to improve the accuracy of the often challenging U2R and R2L classes, which are known weaknesses in many IDS solutions.

Next, they conducted feature extraction in both groups. Initially, they employed Intersectional Correlated Feature Selection (ICFS), which uses the Pearson Correlation Coefficient (PCC) to select significant features between two random variables. PCC quantifies the degree of variation between two variables. Following ICFS, Principal Component Analysis (PCA) was utilized to reduce the data's dimension.

To maintain a 1:1 ratio between attacks and normal data in the second group, an equal amount of data was randomly chosen to match the combined R2L and U2R classes. These two groups were then used to train their model, which consisted of a first layer utilizing the Naïve Bayes classifier and a second layer employing SVM. The first layer's role was to detect DoS and Probe classes, and if the outcome did not belong to these two categories, the data proceeded to the second layer to ascertain whether it belonged to the R2L, U2R, or Normal classes. Their solution was evaluated using the NSL-KDD dataset, achieving an overall accuracy of 93.11%, and demonstrating a detection rate of 96.67% for R2L and 100% for U2R classes. While their solution excelled at identifying small classes, it did not achieve the same level of accuracy for the larger classes, which contrasts with other efficient solutions.

Yiping, et all proposed a Hybrid Nested Genetic-Fuzzy Algorithm (HNGFA) for attack detection. They initiated the process with feature selection using the Naïve Bayes method, which divided features into two groups: Major features and Minor features. Their model comprises two genetic-fuzzy algorithms, specifically the Outer Genetic-Fuzzy Algorithm (OGFA) and the Inner Genetic-Fuzzy Algorithm (IGFA). Each of these algorithms incorporated two nested genetic algorithms. The outer algorithm managed fuzzy sets, while the inner algorithm governed fuzzy rules.

The OGFA was tasked with classifying data with major features, while the IGFA handled data with minor features. These two genetic-fuzzy algorithms interacted to generate new solutions, striving for improved accuracy. The objective was to facilitate collaboration between the best results from the OGFA and the weaker outcomes from the IGFA to create the most effective model. Their solution was evaluated on the KDDcup99 and UNSW-NB15 datasets, resulting in respective overall accuracies of 98.19% and 80.54%. Furthermore, their solution demonstrated a strong accuracy in detecting smaller classes such as R2L and U2R. However, it's important to note that the complexity of their model resulted in longer training times."

Machine Learning

Machine learning is a branch of artificial intelligence (AI) that deals with the creation of statistical models and algorithms that let computer systems learn from data without explicit programming, allowing them to perform better and better over time on a given task. To put it another way, machine learning systems use data to identify trends, anticipate outcomes, or guide decisions. As they learn from experience and are exposed to more data, they improve at what they do. It's a broad field with applications in natural language processing, recommendation systems, picture and speech recognition, and more. Depending on the nature of the learning process and the available data, machine learning comprises several learning styles, including supervised learning, unsupervised learning, and reinforcement learning.

Types of Machine Learning

a) Supervised Learning

- In supervised learning, every example in the training dataset is linked to a target or result, and the model is trained using labelled data.
- Learning a mapping from inputs to outputs is the aim, which makes it applicable to problems such as regression (predicting continuous values) and classification (assigning labels to data points).
- Common algorithms include neural networks, logistic regression, decision trees, support vector machines, and linear regression.

b) Unsupervised Learning

- Unsupervised learning involves finding patterns or structure in unlabelled data, where there are no predefined target labels.
- Clustering, or putting related data points together, and dimensionality reduction, or lowering the amount of features while keeping crucial information, are frequent tasks.
- Common algorithms: K-means clustering, hierarchical clustering, Principal Component Analysis (PCA), and auto encoders.

c) Reinforcement Learning

- Through interactions with its surroundings, an agent that uses reinforcement learning learns how to maximize a cumulative reward signal by performing a series of decisions or actions.
- It is frequently applied to situations like gaming, robotics, and autonomous systems where decisions have long-term effects.
- Common algorithms: Q-Learning, Deep Q- Networks (DQN), and Proximal Policy Optimization (PPO).

Datasets

The researchers used datasets to test and train their models. The most well-known and often used datasets for training and testing intrusion detection systems are covered in the section that follows.

Types of Datasets

a) **KDDcup99**

One of the datasets that is most frequently used to evaluate IDS is KDDcup99. On the DARPA'98 dataset, it is based. The number of samples in the KDDcup99 is roughly 4,900,000. Every sample is labeled as Attack or Normal and has 41 attributes. The attack samples have been categorized into four groups: Probe, User to Root (U2R), Denial of Service (DoS), and Remote to Local (R2L). For KDDcup99, there are three distinct datasets: the entire dataset, a dataset that represents 10% of the entire dataset, and a test dataset with 311,029 samples. This dataset's imbalance, or the fact that many examples for large classes like DoS and Probe are similar to one other but there are few for R2L and U2R, is one of its main drawbacks. Depending on the portion of the dataset that is utilized, certain classes may not exist at all.

b) **Kyoto 2006**

In order to gather different kinds of traffic, Kyoto University's exterior and interior network security measures, email servers, web crawlers, honeypots, and dark net sensors were all used to build this dataset. They extracted 14 statistical features from the KDDcup99 dataset based on its 41 features. Every sample has 24 features because they additionally extracted 10 more features to create the dataset. The Kyoto dataset's most recent version covers traffic from 2006 to 2015.

c) **NSL-KDD**

The KDDcup99 dataset's primary problem was resolved with the creation of this dataset. It was suggested by Tavallaee and associates in 2009. The KDDcup99's four assault categories are retained. Two files—a training set and a testing set—are suggested by the NSL-KDD. There are 126,620 instances in the training set, consisting of 21 distinct attacks. There are 22,850 occurrences in the testing set, which consists of 37 distinct attacks.

d) **UNSW-NB15**

The Australian Centre for Cyber Security produced this dataset. Its purpose was to produce traffic—a cross between regular activity and assault behaviors. There are nine different kinds of assaults in this dataset: worms, shellcode, fizzers, analysis, backdoors, DoS, exploits, and generic attacks. Two files—a training set and a testing set—are suggested by UNSW. These

files include entries from the original dataset pertaining to various traffic kinds, including attacks and regular traffic. There are 2,540,044 records in the original dataset, 175,341 records in the training set, and 82,332 records in the testing set.

e) CICIDS2017

The Canadian Institute for Cybersecurity (CIC) produced this dataset in 2017. Real-world traffic comprising both typical and recent attack samples was used to create this dataset. Using CICFlowMeter, the results were analyzed according to the time stamp, source and destination IP, protocols, and assaults. Furthermore, they employed standard techniques including Distributed Denial of Service (DDoS), Heart Bleed, Web Attack, Infiltration, Botnet, and Brute Force FTP and SSH.

METHODOLOGY

Creating an efficient Intrusion Detection System (IDS) utilizing machine learning demands a systematic methodology. It commences with a precise delineation of the system's goals and prerequisites, which is instrumental in determining the specific threats to be detected and the unique network requirements. The subsequent stage involves the collection and preprocessing of data, where pertinent datasets are acquired and cleansed for the purposes of training and testing. A critical juncture in the process is the selection of the most appropriate machine learning algorithms and feature extraction techniques, enabling the system to discern patterns and anomalies in network traffic.

Machine learning models are then subjected to rigorous training and testing using historical data, enhancing their capability to identify both known and potential unknown threats. Real-time monitoring is implemented to ensure the IDS's effectiveness in recognizing intrusion attempts as they unfold. To swiftly respond to security breaches, alerting mechanisms and response protocols are carefully devised.

The process necessitates ongoing refinement and optimization to stay in step with the perpetually evolving threat landscape, mandating regular updates to the machine learning models. The integration of compliance and reporting requisites ensures that the system aligns with industry standards and regulatory frameworks. Furthermore, maintenance, comprehensive documentation, and continuous staff training are indispensable elements in preserving the IDS's efficacy over time, positioning it as a proactive and dependable cybersecurity defines mechanism.

RESEARCH RESULT AND DISCUSSION

In brief, this paper offers a comprehensive examination of Intrusion Detection Systems (IDS) and their potential improvement through machine

learning integration. It categorizes IDS types, discusses attack detection methods, evaluation metrics, and machine learning techniques. Recent trends show a shift toward deep learning for better attack detection, albeit with increased complexity and resource demands. Feature extraction techniques like Auto-Encoder are also gaining prominence. Recommendations include further deep learning optimization, efficient resource allocation, interdisciplinary collaboration, user education, and regular maintenance. Research and innovation in IDS, particularly with machine learning, are vital for adapting to the evolving cybersecurity landscape and safeguarding network integrity and data security.

Future Trends

Future trends in Intrusion Detection Systems (IDS) include improved detection of zero-day attacks, enhanced AI and machine learning capabilities, increased use of cloud-based solutions, a focus on behavioural analysis, integration of threat intelligence, automated response mechanisms, and user and entity behaviour analytics. Additionally, IDS will adopt privacy-preserving techniques, blockchain-based logging, and homomorphic encryption, while also providing customization and fine- tuning options to organizations. IDS will continue to play a pivotal role in regulatory compliance efforts, underscoring the dynamic nature of the cybersecurity landscape and the need for innovative approaches to address evolving and sophisticated threats.

CONCLUSIONS

In summary, this paper provides a comprehensive exploration of Intrusion Detection Systems (IDS) and their potential for improvement through machine learning integration. It begins by elucidating the core concepts of IDS, categorizing them into Network Intrusion Detection Systems, Host Intrusion Detection Systems, and Hybrid Intrusion Detection Systems, while also addressing their methods of attack detection, involving both recorded signatures and behavioural analysis against a baseline of normal network activity. The paper further delves into the crucial metrics used to evaluate IDS, focusing on key measures such as Accuracy, Detection Rate (Recall), and F-Measure. A broad overview of machine learning is provided, categorizing techniques into three primary types: Supervised, Semi-supervised, and Unsupervised learning. Moreover, a comprehensive review of recent research papers utilizing machine learning for IDS is presented, with a noticeable shift towards deep learning methods for enhanced attack detection, albeit with increased model complexity and computational demands. Additionally, the growing trend of feature extraction methods, including Auto-Encoder, is highlighted.

In light of these findings, several recommendations emerge: the need for continued exploration and optimization of deep learning models in IDS, efficient resource allocation for supporting computationally intensive models, interdisciplinary collaboration between cybersecurity and machine learning experts, user education to bolster security awareness, and regular maintenance

to adapt to evolving threats. Given the ever- evolving cybersecurity landscape, research and innovation in the field of IDS, particularly in conjunction with machine learning, are of paramount importance to effectively safeguard network integrity and data security.

ADVANCED RESEARCH

Considering the researchers' own limited knowledge and skills, the researcher has come to the realization while producing this article that there are still numerous deficiencies in language, writing, and presentation style. As a result, the researcher anticipates helpful critiques and recommendations from a range of sources to ensure the piece is flawless.

ACKNOWLEDGMENT

Any project's ability to succeed is mostly dependent on the support and direction of numerous other people. Without their assistance, this research effort would not have been feasible. We would want to take this opportunity to thank everyone who has contributed to the successful completion of this project. First and foremost, we would like to express our profound thanks to our respected HOD, Mrs. Meenakshi Thalor, and mentor, Ms. A.S. Phapale, for their ongoing support and encouragement during the preparation of this report, as well as for making the library resources necessary for this preparation available. Our many conversations were very beneficial. We owe her a great deal for her direction, persistent watchfulness, provision of pertinent course information, and assistance in seeing the seminar through to completion. We respect her because of the leadership, inspiration, and encouragement she has given us.

REFERENCES

Anderson, P. Computer Security Threat Monitoring and Surveillance. 1980. Available online: https://csrc.nist.gov/csrc/media/ publications/conference- paper/1998/10/08/proceedingsof-the-21st- nissc-1998/documents/early-cs- papers/ande80.pdf (accessed on 19 May 2022.

Checkpoint. What Is an Intrusion Detection System? Available online: https://www.checkpoint.com/cyber-hub/networksecurity/what-is-an-intrusion- detection-systemids/ (accessed on 19 May 2022).

IBM Cloud Education. Unsupervised Learning. Available online: https://www.ibm.com/cloud/learn/unsupervi sed-learning (accessed on 19 May 2022)

IBM Cloud Education. Machine Learning. Available online: https://www.ibm.com/cloud/learn/machine- learning (accessed on 19 May 2022).

IBM Cloud Education. Supervised Learning. Available online: https://www.ibm.com/cloud/learn/supervise d-learning (accessed on 19 May 2022).

Sabahi, F.; Movaghar, A. Intrusion Detection: A Survey. In Proceedings of the 2008 Third International Conference on Systems and Networks Communications, Sliema, Malta, 26–31 October 2008; pp. 23–

Seldon. Machine Learning Regression Explained. Available online: https://www.seldon.io/machine- learning-regressionexplained (accessed on 19 May 2022).

Terence, S. All Machine Learning Models Explained in 6 Minutes. Available online: https://www.ibm.com/cloud/learn/ unsupervised-learning (accessed on 19 May 2022).

ThreatStack. The History of Intrusion Detection Systems (IDS)—Part 1. Available online: https://www.threatstack.com/blog/ the-history-of-intrusion-detection-systems- ids-part-1 (accessed on 19 May 2022).