A review of Artificial Intelligence Algorithms (Machine Learning Algorithm) for Intrusion Detection in Software-Defined Networking

Reelyn Heymann¹ and Victor Ahirhima¹

¹University of Johannesburg

March 12, 2024

Abstract

Demands for flexible and seamless system management necessitated the growth of software-defined networks (SDN). Yet, securing these environments with effective measures is critical as SDN continues to evolve into more intricate architectures. Intrusion detection is paramount among these measures; thus far, studies suggest that artificial intelligence (AI) approaches may be helpful in this domain. By systematically examining relevant works addressing AI-based intrusion prevention strategies within hyper-evolved SDN settings, our review aims to present an inclusive evaluation alongside suggesting areas requiring additional scrutiny. This research introduces readers to key concepts related to SDN and how deep learning algorithms, machine learning algorithms, and neural networks can be applied for effective intrusion detection within an SDN environment. Drawing from existing literature on this subject matter, our analysis critically examines the benefits and drawbacks of these AI-based techniques while highlighting gaps in knowledge requiring further research attention. Some areas include real-time protection capabilities, scalability concerns, and seamless integration with different security mechanisms. We then present future research directions in this area. This literature review employs a systematic approach to elucidate the current research on using AI methods to detect intrusions in SDN.
Manuscript Title:


Author’s:

Heymann R.

Ahirhima V.

Submitting Author’s Email address:

rheymann@uj.ac.za {Insitutional}

reolyn.heymann@gmail.com {Personal}

In text Citation for figures:

Figure 1: A Block Diagram of IDS [15].


Figure 2: The Block Diagram of SDN Architecture [27].

1. INTRODUCTION

As network security concerns grow in sophistication each day, Artificial Intelligence (AI) represents a beacon of hope in combating against emerging threats emanating from varied sources. One key area where AI has proven effective is detecting potential threats to Software Defined Networking (SDN). This particular framework employs the utilization of the division between a network control plane and a data plane to provide a centralized means of control over diverse traffic dynamics and maximize flexibility. Still, it presents unique risks requiring the implementation of Intrusion Detection (ID) systems. These systems are pivotal in securing networks employing both conventional and SDN architectures by monitoring traffic behavior for anomalous activity that might signal intrusion attempts or attacks on the integrity of the network.

In traditional settings, ID solutions employ rule-based or signature-based detection algorithms crafted explicitly for known attack scenarios; however, they might not be as effective given their narrow scope of operation under such circumstances when tasked with securing dynamic settings such as SDN networks. Before the rapid growth of attacks and increased global traffic, signature-based Intrusion Detection Systems (IDSs) have traditionally served as the industry standard for network intrusion detection. Nonetheless, such systems remain limited to detecting only known attacks and are plagued by issues related to scalability. It is worth knowing that the availability of acceptable datasets for training IDSs remains a challenge, compounded by the fact that many such datasets have been subject to intense critique regarding their accuracy and ability to mirror real-world settings. In light of these challenges, the use of Machine Learning has emerged as a promising avenue for the development of IDSs [1].

The 2017 Internet Security Threat Report by Symantec [63], has exposed that the magnitude and seriousness of zero-day attacks in 2016 were vastly more significant than in previous years, with an unbelievable three billion attacks reported. As per the statistics about data breaches, hackers have been held responsible for the loss or theft of nearly nine billion data records since 2013. While cybercriminals previously focused on stealing credit card details or robbing customers' bank accounts, the latest malware is more audacious and targets banks directly, with some attacks attempting to steal millions of dollars in one go.
Additionally, Symantec has witnessed a rise in the instances of security breaches. Consequently, detecting zero-day attacks has become the utmost priority [2]. Machine learning has made great strides in various areas, and one notable field is Cybersecurity. The remarkable capability of deep learning algorithms to detect previously unknown zero-day attacks has profoundly impacted the realm of Cybersecurity. A "Day Zero attack" refers to a type of cyberattack that takes advantage of a previously unknown vulnerability in a software program, operating system, or device. This vulnerability is called "zero-day" because the developers have had zero days to fix or patch it before attackers start exploiting it [64]. These advanced algorithms have been extensively utilized across various cybersecurity tasks, from analyzing malware to identifying unauthorized intrusions and malicious botnets.

By harnessing the power of deep learning, significant enhancements have been achieved in these areas, ultimately bolstering the security and safeguarding of our online environment [3]. The fundamental objective of this paper is to furnish a complete and all-encompassing evaluation of the literature concerning the application of AI approaches to identify breaches in the context of software-defined networking. This paper intends to examine and assess the current state of understanding in this domain, pinpoint any gaps or restrictions in the existing investigation, and propose potential pathways for future research endeavors to tackle those limitations. A systematic methodology was employed to search, assess, and scrutinize pertinent literature. Initially, specific keywords such as "intrusion detection," "machine learning," "software-defined networking," "cybersecurity," and "artificial intelligence" were utilized to retrieve articles from scholarly journals, conference proceedings, and technical reports. Afterward, every piece was methodically assessed based on its importance, design excellence, and writing excellence, focusing on methodology, outcomes, and conclusions. Upon identifying relevant articles, they were juxtaposed to identify patterns and trends in the literature. Ultimately, a comprehensive summary was composed, encapsulating the most important information from each study while ensuring accurate and complete source tracking. Overall, a precise and methodical technique was implemented to investigate the literature on the main use of AI for intrusion detection in Software-Defined Network, resulting in a thorough knowledge of the existing research on the subject and a significant contribution to the field. The subsequent parts of this review paper are organized in the following manner. In Section 2, we shall provide a concise explanation of the fundamental knowledge. Section 3 is dedicated to presenting a review of case studies and
relevant investigations. In Section 4, we examine an analysis of the numerous evaluation metrics and performance assessments of AI-based Intrusion Detection Systems in Software-Defined Networking. Section 5 delves into the challenges and suggests future avenues of research. Finally, Section 6 concludes our work.

2. BACKGROUND

In the background section of our review paper, we provide a concise overview of the fundamental concepts and technologies relevant to our research on utilizing artificial intelligence (AI) techniques for intrusion detection in Software Defined Network (SDN). An initiatory section on employing machine learning algorithms and deep learning algorithms in intrusion detection and a talk about the diverse sorts of intrusion detection systems are comprised. Furthermore, an overview of SDN, its architecture, and its distinctive security challenges are presented. Finally, the potential benefits of AI techniques in SDN intrusion detection are emphasized. This background information contextualizes the significance of our research.

2.1 Machine Learning Algorithm in Intrusion Detection

Machine learning techniques have emerged as vital tools in bolstering our defenses against attacks. These technologies, grounded in machine learning principles, can accurately differentiate between regular network traffic and malicious attacks. Numerous research initiatives have been dedicated to developing effective methods for detecting and preventing various threats arising in Software-Defined Network (SDN). These endeavors do aim to enhance the security posture and safeguard against potential vulnerabilities within SDN environments [4]. Various approaches based on machine learning were utilized to address the challenges of network intrusion detection. These approaches typically consist of three key phases. Firstly, in the preprocessing phase, the collected data instances from the network environment are organized and prepared for input into the machine learning algorithm. This phase also involves extracting relevant features from the data and selecting the most informative ones. Secondly, a machine learning algorithm is employed during training to learn and characterize the patterns inherent in different data types. By doing so, it constructs a system model that captures the cardinal structures and characteristics of the data. Lastly, in the detection phase, the established system model is utilized to evaluate incoming traffic data. This involves comparing the observed patterns in the data with those stored in the system model. If a match is found indicating a potential threat, an alarm is triggered
to notify the appropriate entities. The three phases collectively form the foundation of machine learning-based network intrusion detection, enabling the detection and identification of potential threats in network traffic data [5]. Machine learning algorithms are leveraged as valuable tools for designing intrusion detection systems (IDS) that target anomalies. The primary challenge lies in crafting a model that not only exhibits high accuracy but also minimizes the occurrence of false positives. A diverse array of algorithmic categories is employed in machine learning to tackle this objective [6]. These categories of ML algorithms encapsulate the range of approaches utilized within machine learning.

1. Supervised Learning: This is a classification approach where training data instances are labeled. It encompasses a range of algorithms such as Artificial Neural Networks, Bayesian Statistics, Gaussian Process Regression, and more. Popular methods include Decision Trees Method, The Support Vector Machines, and Ensemble classifiers. These algorithms are vital in analyzing data and making predictions based on labeled examples, enabling accurate classification in various applications [7]. Supervised learning, on the other hand, leverages labeled data to train models that can classify network traffic as usual or malicious. By learning from known examples of attacks or benign behavior, these models can accurately identify and categorize new instances, helping detect and prevent potential security breaches.

2. Unsupervised Learning: The algorithms are designed to grasp the underlying patterns and representations within unlabeled input data in an unsupervised learning approach. These algorithms aim to capture the fundamental structure or distribution in the data, enabling them to make predictions on new, unknown data. Unsupervised learning encompasses various techniques such as principal component analysis (PCA), which aims to reduce the dimensionality of the features, and clustering techniques like the self-organizing map (SOM), which groups similar instances together [8]. Unsupervised learning allows us to discover patterns and anomalies in network data without relying on pre-labeled examples. By analyzing the behavior of network traffic or system logs, we can uncover hidden threats or unusual activities that may indicate an intrusion.
3. Semi-Supervised Learning: This type of supervised learning involves using labeled and unlabeled data for training. Instead of relying solely on labeled examples, this approach incorporates a significant amount of unlabeled data into the learning process [8]. Semi-supervised learning offers a practical and highly effective approach to intrusion detection. Combining labeled and unlabeled data enables us to improve the accuracy and efficiency of our detection models. This is particularly valuable when we have limited or scarce labeled data available.

4. Reinforcement Learning: This constitutes a conceptual framework that entails an agent, a state space (S), and an action space (A). The agent, as an entity aimed at learning, engages with its environment in order to ascertain the optimal action to maximize the long-term reward. This long-term reward combines immediate and future rewards, where the future rewards are discounted to account for their delayed impact. In simpler terms, RL allows the agent to acquire knowledge from its environment over time and make decisions that lead to the best cumulative rewards, considering the immediate gains and the potential benefits in the future [9]. Reinforcement learning provides an exciting perspective for intrusion detection. By viewing the agent as the defender and the environment as the network, we can train the agent to make optimal decisions based on feedback received as penalties or rewards. This approach enables the agent to learn from past experiences, adapt strategies, and effectively respond to evolving attack techniques.

2.2 Deep Learning Algorithm in Intrusion Detection

Deep learning algorithms have shown their effectiveness in intrusion detection systems. They use artificial neural networks to automatically learn and recognize patterns that suggest malicious activities or intrusions in a network. These algorithms can identify abnormal behaviors and potential threats by analyzing network data without relying on predefined rules. There are various Deep Learning Algorithm which will be discussed below:

1. Convolutional Neural Network (CNN): This type of neural network specializes in processing and categorizing image data. One of its key strengths is its ability to extract relevant information and simplify complex features, which makes it highly applicable in
Convolutional Neural Networks (CNN) can be instrumental in detecting intrusions within a network. It analyzes network traffic data and identifies unusual patterns indicating potential unauthorized access or malicious activities. By incorporating CNN into intrusion detection systems, network administrators and security professionals can enhance their ability to detect and respond to potential threats swiftly.

2. Recurrent Neural Network (RNN): This a neural network that builds on the feed-forward model by incorporating sequential information. The RNNs are deemed recurrent as they execute identical tasks for each element within a sequence, where the output is reliant on the antecedent computations. RNNs are a type of deep learning algorithm that excels at processing sequential data, making them well-suited for analyzing network traffic patterns over time.

3. Long Short-Term Memory Networks (LSTMs): The LSTM model is highly appropriate for handling and forecasting crucial occurrences characterized by prolonged intervals and temporal lags in time series data. In contrast, conventional RNN networks rely on the back-propagation through time algorithm for training. Nevertheless, when the duration of time is significantly extended, the residual error that necessitates return decreases exponentially, thus leading to a tardy update of network weight. This sluggishness cannot adequately reflect the impact of the RNN's long-term memory. LSTMs learn from a large set of standard network behavior data, allowing them to understand what is considered "normal." When faced with new network traffic data, they can compare it to their learned model and identify any deviations or anomalies that might indicate an intrusion.

4. Generative Adversarial Networks (GANs): This type of algorithm can generate synthetic data that is realistic and authentic. By training the generator aspect of a GAN on copious amounts of actual network traffic data, it can fabricate synthetic network data that closely emulates actual network behavior.
5. Autoencoders: These are neural networks commonly used for unsupervised learning and data compression. They have two main components: an encoder and a decoder. Autoencoders are also effective in anomaly detection, as they learn to reconstruct normal patterns and can detect deviations or abnormalities in the input data. Additionally, autoencoders are valuable for dimensionality reduction, where they help in reducing the complexity and size of high-dimensional datasets while preserving the most important information [14]. Zhu et al. in [14] utilized an autoencoder, a formidable neural network architecture, in their scholarly inquiry. The purpose behind utilizing the autoencoder was to acquire meaningful and condensed features from the 2D images obtained by projecting 3D shapes into a 2D space. Through training the autoencoder on these 2D images, the model gained the ability to assimilate crucial information about the shapes and represent them in a concise manner.

2.3 Intrusion Detection System (IDS)

The key objective of the IDS is to detect and alert system administrators of any suspicious or unauthorized behavior that may indicate an ongoing cyber-attack. AI-powered IDS systems can examine network traffic in real-time, detect abnormalities or anomalous behavior, and create alerts or execute automated measures to mitigate potential threats.

We primarily have two types of IDS, which could complement each other in an ideal setting. They include:

1. Network-based IDS (NIDS): This Intrusion Detection System (IDS) uses live network traffic monitoring to detect malicious activity. Specifically, it scrutinizes the data packets coursing through the network to identify any recognizable attack patterns or anomalous behavior. Positioned strategically at pivotal points within the network infrastructure, such as switches or routers, the Network-based Intrusion Detection System (NIDS) captures and thoroughly inspects the traffic. With its broad perspective on network security, NIDS is highly effective at thwarting assaults that target multiple systems [16].

2. Host-based IDS (HIDS) operates on individual host machines to observe and scrutinize each host's activities. It analyzes various factors such as system logs, file alterations, and
the registry to identify suspicious behavior. HIDS can uncover attacks that may otherwise go unnoticed by network-based IDS, including endeavors to achieve unauthorized access or elevate privileges on a specific host. While HIDS provides a more comprehensive level of security, it necessitates deployment and administration on each host [17].

Traditional intrusion detection systems (IDS), comprising network-based IDS (NIDS) and host-based IDS (HIDS), confront specific challenges and possess certain limitations. One of these limitations is the potential for producing false positive alerts, which can result in unnecessary inquiries and depletion of resources [18]. The traditional IDS predominantly depends on established attack signatures or predefined rules, which can create difficulty in identifying novel or unknown types of attacks, leaving systems susceptible to emerging threats [19]. Moreover, the traditional IDS may encounter scalability issues while dealing with large-scale networks or environments with numerous hosts. Managing and deploying IDS on a large scale can be intricate and resource intensive. AI-based IDS systems have the potential to overcome these challenges and significantly enhance the effectiveness and efficiency of intrusion detection and response [20].

2.4 Software-Defined Network

Software-Defined Network (SDN) is a contemporary methodology of network architecture that endeavors to streamline and amalgamate the administration and supervision of network infrastructure. In customary networking, network devices, for instance, switches and routers, possess embedded control functions that ascertain how traffic circulates within the network. Conversely, SDN bifurcates the control plane (which affirms decisions regarding traffic handling) from the data plane (which essentially forwards the traffic) [21].

The SDN architecture comprises three fundamental components: the application layer, the control layer, and the infrastructure layer.

Various network services and applications are deployed at the SDN architecture's application layer. Network applications can be developed by network administrators, third-party vendors, or end-users themselves within the SDN architecture's application layer. By leveraging the SDN architecture's programmability, they can execute network policies, traffic management, security measures, and other functionalities [22].
The control layer involves the software controller, which acts as the central point of control for the network. The software controller at the control layer interacts with applications deployed at the application layer and communicates with network devices in the infrastructure layer. The data gathered regarding network topology, traffic patterns, and network policies is utilized by the SDN architecture's software controller at the control layer to make decisions on handling and routing network traffic. The controller then transmits these decisions to the network devices, which execute the actual forwarding of the traffic [23].

The layer of infrastructure is constituted of physical network devices, including switches, routers, and other types of networking hardware. The forwarding of network traffic is executed by the physical network devices in the infrastructure layer based on the instructions received from the controller. In SDN architecture, the network devices have a simplified role as they primarily focus on packet forwarding, while the control and intelligence are centralized in the software controller [24].

When considering Software-Defined Networking (SDN), it is imperative to address several security challenges. A primary concern pertains to the centralized nature of SDN controllers. These controllers determine the forwarding of data packets within the network and, as such, are vulnerable targets for malicious actors. Should these controllers be compromised, unauthorized access, data breaches, and consequential disruptions to the entire network are likely to occur [25].

By meticulously gathering and scrutinizing data from various sources, including network devices, endpoints, and security logs, the integration of AI substantially improves network visibility. This all-embracing perspective provides a complete approach to identifying potential security gaps, vulnerabilities, and misconfigurations that might otherwise go unnoticed [26].
3. REVIEWS OF RELATED WORK

This section scrutinizes extant literature on the part played by Artificial Intelligence (AI) techniques in detecting intrusions within Software-Defined Networking (SDN). The contributions and methodologies employed by prior researchers in this domain are thoroughly deliberated.

In [28], the author's primary focus was on the control plane of SDN, and they employed anomaly detection methods to enhance Distributed Denial of Service (DDoS) attack detection capabilities. While acknowledging the existence of prior works that proposed various techniques for DDoS attack detection within SDN layers, the author underscored the necessity for further advancement in this arena. To summarize, the author's objective was to refine DDoS attack detection mechanisms on the control plane of SDN by capitalizing on anomaly detection techniques. Machine learning algorithm was utilized for anomaly detection.

In [29], the authors propose an Intrusion Detection System (IDS) that utilizes a GridSearch technique with Support Vector Machine (SVM) for identifying anomalies associated with attacks. The findings indicate significant progress in identifying various network attacks within an SDN-based cloud environment, highlighting the promising potential of the proposed system to enhance security measures in such settings. According to the findings, the proposed IDS displayed a remarkable detection rate, achieving an accuracy of over 99.8 percent with the implemented machine-learning model. These outcomes signify noteworthy advancements in identifying various network attacks within an SDN-based cloud environment. The proposed system demonstrated proficiency in identifying and categorizing almost all possible network attacks, thereby highlighting its promising potential to strengthen security measures in such settings.

In [30], the authors investigated SDN-based detection systems designed for identifying DDoS attacks by employing machine learning with the help of the feature selection method. The suggested methodologies manifested remarkable accuracy in detection, where the initial approach achieved 98.3% accuracy without discriminating against traffic types, and the subsequent process achieved 97.7% sensitivity by categorizing DDoS attacks and alleviating the burden on the controller.
In [27], the authors presented a compelling concept for detecting malicious activities in the SDN data plane. They emphasized the benefits of employing MADMAS (Malicious Activity Detection in the SDN Data Plane) compared to alternative solutions. However, they recognized the necessity for further improvements, particularly in efficiently identifying U2R attacks, which may require incorporating deep packet inspection techniques. Additionally, the article explored the efficacy of machine learning methodologies such as Self Organizing Maps and Learning Vector Quantization, in conjunction with their enhanced versions, and juxtaposed the outcomes with other SDN-based Intrusion Detection Systems (IDS).

In [31], the authors highlight the advantages of integrating blockchain, SDN, and CIDS (Collaborative Intrusion Detection System) to propose a comprehensive framework named BlockCSDN for collaborative intrusion detection in SDN. The authors have implemented an implementation instance that employs a challenge-based CIDS and has conducted experiments in simulated and real network environments to assess its performance. The experiments' outcomes substantiate the approach's feasibility and efficiency in addressing insider attacks, preserving network bandwidth, and enhancing the resilience of alarm aggregation.

In [32], extensive experimentation was conducted using three benchmark datasets, NSL-KDD, UNSW-NB15, and AWID, to assess the proposed model's performance. The model demonstrated superior accuracy and a lower False Positive Rate (FPR) compared to the existing state-of-the-art systems. Furthermore, the model's robustness against adversarial attacks was analyzed, revealing only a marginal decline in accuracy compared to other models. To further enhance the system's robustness, the concept of denoising autoencoder was implemented. The system's usability in real-life scenarios was also demonstrated, showcasing its adaptability to changes in attack patterns.

The authors in [10] draw attention to a noteworthy transition in research emphasis within the domain of Intrusion Detection Systems (IDS), moving away from supervised learning and towards clustering and alternative algorithms. This shift holds much promise for the future as it indicates that IDS will become more adept at detecting unknown and zero-day attacks. The study additionally observes a substantial rise in deploying hybrid algorithms from 2019 to 2022, in contrast to 2015-2018. Furthermore, there is an emerging tendency to create modern IDS systems
utilizing more recent datasets. Interestingly, the study highlights that several machine learning algorithms remain unexplored in IDS design, suggesting potential avenues for future research.

In [33], the authors present a novel methodology dubbed NIDS-DL (Network Intrusion Detection System utilizing Deep Learning) that is specifically tailored for Software-Defined Networking (SDN). The efficacy of various deep learning algorithms was assessed and evaluated using an array of metrics, such as Accuracy, F-score, Recall, and Precision. Feature selection methods were employed to train the classifiers using highly correlated features to improve performance further. The proposed approach was then applied to the widely recognized NSL-KDD dataset, commonly adopted for network intrusion detection research.

A novel method is proposed in [34] that amalgamates statistical and machine learning techniques. At the outset, a correlation-based strategy with a dynamic threshold was implemented. However, the outcomes were not deemed satisfactory. The experiments on multiple datasets with high False Positive Rates (FPR) revealed this inadequacy. To overcome this constraint, a range of machine learning algorithms were employed. The investigation results indicate that the model proposed in this study outperforms other models in accuracy and precision.

The authors in [34] propose a statistical methodology to address the subpar results of traditional intrusion detection techniques. The proposed methodology employs a multilayer convolutional neural network for the purpose of feature extraction and selection, and a softmax classifier for intrusion classification. In addition to this, a multilayer deep neural network is employed to facilitate further analysis and variety. The efficacy of this approach is evaluated using two widely used benchmark datasets, namely NSL-KDD and KDDCUP'99. The assessment is carried out using performance metrics such as accuracy, recall, F1-score, and precision. The experimental results reveal that the proposed approach achieves a superior accuracy rate of 99% when compared to other Intrusion Detection Systems.

The authors of this study in [35] proposed an intrusion detection model incorporating an attention mechanism. The model is comprised of a convolutional neural network, a denoising autoencoder, and an attention mechanism. The CNN retrieves abstract characteristics while the DAE eliminates irrelevant features to enhance the model’s efficacy in subsequent training. The attention mechanism prioritizes important features by assigning them greater weight, amplifying
their influence on the model. The authors utilized the Jupyter Notebook platform and the Anaconda environment to construct and train the model. The model was evaluated using the In SDN dataset, which is exclusive to the SDN environment, and compared with the performance of CNN, LSTM, and CNN+LSTM models. The DCA (DAE-CNN-Attention) model outperformed other models by 0.6%, achieving an accuracy of 99.7% during binary classification testing.

Kavin B et al. in [36] engaged in a comparative evaluation of the pre-existing systems with the objective of evaluating their effectiveness. Subsequent to this assessment, some proposed innovative concepts and methodologies that are specifically targeted at enhancing the efficacy of the extant intrusion detection systems was explored. These proposals may involve improvements to the detection algorithms, data preprocessing techniques or the overall system architectures.

Alatwi et al. in [37], proposed a study which furnishes a thorough and all-inclusive analysis of black-box adversarial attacks, which are aimed specifically at Machine Learning (ML)-based Network Intrusion Detection Systems (NIDS). The survey comprises an extensive review of the literature, addressing diverse attack techniques and strategies that are utilized to mislead ML-based NIDS.

Ahmed et al. in [16], presented a unique method that integrates an active learning model based on entropy for the efficient detection of intrusion patterns. The proposed model utilizes the idea of entropy uncertainty to make informed decisions regarding which instances should be included in the training set. Through the implementation of a pooling strategy, the model can selectively add instances to the training set, focusing primarily on those that exhibit higher levels of entropy uncertainty. This approach is intended to enhance the decision boundary of the model by incorporating instances that are more difficult to classify or possess greater information gain.

Shinde et al. in [26], presented a study which directs its attention towards the seamless incorporation of Quantum Machine Learning (QML) and Software Defined Network (SDN) technologies, with the objective of investigating their potential advantages and interactions. The SDN paradigm is widely acknowledged as a profound advancement over conventional network architectures, and Machine Learning and Artificial Intelligence have been instrumental in shaping numerous facets of SDN.
Kurochkin et al. in [18], have conducted a comparative analysis of machine learning methods utilizing neural networks on a carefully selected dataset. The outcomes of their experimentation and analysis have conclusively shown that the application of deep neural networks can effectively tackle the challenge of intrusion detection in software-defined networks. In light of these results, it can be inferred that deep neural networks hold immense potential in accurately detecting and categorizing various network attacks in the domain of software-defined networks.

Bawany et al. in [38], gave a noteworthy contribution in two pivotal domains. Firstly, a thorough and perceptive analysis of the subject of DDoS attack detection and mitigation mechanisms, in the context of the Software Defined Networking (SDN) paradigm. The survey comprehensively encompasses diverse techniques that are applied for detecting DDoS attacks and further categorizes these techniques based on their respective detection methodologies. This exhaustive review serves as a valuable repository of knowledge, providing a comprehensive understanding of the contemporary landscape of SDN-based DDoS defense mechanisms.

O'meara et al. in [1], have successfully established elevated baseline scores with respect to key metrics, thereby underscoring the importance of False Negative Rate (FNR) as a critical indicator of model performance. Moreover, they have acknowledged the pressing need to effectively detect hitherto unseen attacks, which is necessitated by the constantly-evolving nature of malicious traffic. In order to evaluate the models' efficacy in detecting such attacks, a holdout testing strategy has been adopted.

Chen et al. in [39] proposed the implementation of this collaborative intrusion detection system denotes a pioneering approach towards augmenting network security within SDN architectures. Through harnessing distributed computational power and empowering switches to operate as neurons, the system proffers a decentralized and scalable resolution for identifying and addressing intrusions.

Amarudin et al. in [40] performed some findings and analysis of the results obtained, it is apparent that a consistently reliable solution for overcoming the challenges associated with intrusion detection has not yet emerged. This observation underscores the need for further development and refinement of intrusion detection models, which can be achieved by leveraging more accurate and reliable classification algorithms.
Finogeev et al. in [25] concentrated on the formulation of an intelligent infrastructure for the transportation environment of the Internet of Things (IoT). The architecture is specifically designed to utilize software-defined network (SDN) and blockchain technologies for the prevention of attacks and detection of threats. The proposed framework has the objective of providing a monitoring system for significant events in the road transport infrastructure.

In this study, blockchain technology plays a critical role in the authentication of network nodes and ensuring the integrity and immutability of sensor data. By exploiting the distributed ledger capabilities of blockchain, the proposed architecture provides a secure and transparent environment for storing and accessing sensor data, thereby augmenting the reliability of the system.

Salih et al. in [41], conducted a study that was focused on the evaluation of the performance of different classifiers that are used in intrusion detection systems. Various metric measurements were applied to assess the classifiers' performance. The study established that the random forest algorithm was able to achieve satisfactory results with the highest accuracy in the classification of different types of attacks.

Moreover, the researcher observed that hybrid classification algorithms were often preferred by researchers when building intrusion detection systems as opposed to individual classification algorithms. This approach involves combining multiple classifiers to leverage their strengths and improve overall performance.

Costa et al. In [42], performed a study which is centered on the provision of a thorough and current examination of the literature concerning the utilization of Machine Learning Techniques in the realm of Internet of Things (IoT) and Intrusion Detection for computer network security. The fundamental aim of this work is to undertake extensive and up-to-date research on pertinent investigations that scrutinize diverse intelligent techniques and their implementation in intrusion detection frameworks within computer networks, with specific attention paid to the Internet of Things and machine learning.

Guo et al. In [43], the scholars introduce three optimization algorithms that are specifically tailored for the operator's network. These algorithms have a focused approach in addressing objectives that are related to network congestion control and prevention, resource preemption,
and balancing network traffic. The goal is to optimize the overall performance of the operator's network. In order to achieve intelligent traffic optimization and enhance network control mechanisms, the scholars have designed optimization algorithms that are customized for each of the objectives. These algorithms utilize the combined capabilities of Software-Defined Networking (SDN) and artificial intelligence. To validate the efficiency of the proposed control mechanism and algorithms, the scholars have conducted large-scale experiments. The results of these experiments demonstrate that the integration of SDN and artificial intelligence in operator networks enables intelligent network control and more efficient traffic optimization.

Adesina et al. in [44], conducted a study with the primary objective of conducting a comprehensive review and analysis of preceding research and efforts within the domain of Network Intrusion Detection. This study serves as a valuable resource for practitioners and researchers engaged in this interdisciplinary pursuit. The authors explored various approaches, algorithms, and techniques that have been developed and applied in the context of Network Intrusion Detection. Furthermore, they critically assessed the strengths and limitations of these approaches, identifying unresolved challenges in the field. My research diverges from the study conducted by Adesina et al. [44] by focusing on the application of artificial intelligence methods for intrusion detection in the realm of Software-Defined Networking (SDN). Unlike their broader review and analysis, my work delves into the intricacies of utilizing AI techniques within the unique context of SDN, contributing to the advancement of intrusion detection solutions tailored to this evolving technology.

The authors in [44] further highlighted recent advancements in hybrid and ensemble systems and provided an overview of existing systems. They carried out an examination of the effectiveness of these systems and identified areas where further improvements can be made.
4. PERFORMANCE ANALYSIS

The present study concerns itself with the performance analysis of AI-based Intrusion Detection Systems (IDS) in Software-Defined Networking (SDN) environments in a human-friendly manner. Performance analysis is a crucial aspect of IDS research, as it involves an in-depth examination of various elements of the IDS to understand how well it performs under real-world scenarios. This study outlines several key performance analysis areas: Detection Accuracy, Scalability, Real-Time Processing, Robustness, Resource Utilization, and Comparative Analysis.

1. Detection Accuracy is a crucial area of performance analysis that examines the IDS's ability to detect and classify different types of attacks accurately. The IDS's ability to distinguish between malicious and benign network traffic is assessed to determine its accuracy [49].

2. Scalability, however, refers to the IDS's performance and efficiency as the network size and traffic volume increase. The IDS's ability to handle more extensive networks and increase traffic loads without compromising its detection capabilities is assessed to determine its scalability [50].

3. Real-Time Processing is another performance analysis area crucial for IDS in SDN environments. The IDS should be able to analyze network traffic and detect attacks in real time without significant delays. The IDS's ability to process traffic quickly and effectively is evaluated to determine its real-time processing capability [51].

4. Robustness is an essential aspect of performance analysis that refers to the IDS's resilience against evasion techniques or attacks designed to evade detection. The IDS's ability to detect and handle such evasion attempts is tested to ensure that it remains effective in challenging scenarios [52].

5. Resource Utilization analysis is another crucial area of performance analysis that assesses the impact of the IDS on network resources such as CPU, memory, and bandwidth. The IDS's efficient operation without consuming excessive resources that could impact the overall network performance is evaluated to determine its resource utilization [37].

6. Comparative Analysis involves comparing the performance of AI-based IDS with traditional approaches or other machine learning algorithms. This Analysis helps
determine whether the AI-based IDS offers improved accuracy, speed, or resource utilization compared to existing solution [53].

By conducting a comprehensive performance analysis, researchers and practitioners can gain insights into the effectiveness, efficiency, and limitations of AI-based IDS in SDN environments. This Analysis helps in identifying areas for improvement, optimizing system parameters, and enhancing the overall performance of the IDS in real-life scenarios.

4.2 Dataset

Selecting a dataset is pivotal when evaluating and analyzing the performance of AI-based Intrusion Detection Systems (IDS) in Software-Defined Networking (SDN). It is imperative that the dataset chosen is representative of actual network traffic and comprises a diverse range of standard and attack instances.

Commonly, researchers opt for benchmark datasets prevalent in the network security and intrusion detection field. Some well-known datasets include the NSL-KDD, an improved version of the popular KDD Cup 1999 dataset. It encompasses various types of attacks and network traffic data generated in a simulated environment.[54] The NSL-KDD dataset serves as a standard benchmark for evaluating IDS performance. Another widely used dataset is the UNSW-NB15, which captures authentic network traffic from a university campus network. It includes both traditional and attack traffic instances and covers many attack types. The AWID (AISense Wireless IDS) dataset targets wireless network intrusion detection. It contains wireless network traffic data, including regular activities and various attacks like DoS and probing attacks.[55] Typically, these datasets are labeled to indicate if each instance represents normal traffic or a specific attack type. Researchers capitalize on these datasets to train and evaluate their AI-based IDS models. Researchers divide the dataset into training, validation, and testing sets during the evaluation process. The training set is utilized to train the AI model, while the validation set assists in tuning the model's hyperparameters. Finally, the testing set is employed to evaluate the model's performance on unseen data [56].

By leveraging these benchmark datasets, researchers can ensure the reproducibility of results and facilitate meaningful comparisons with other IDS models. Nevertheless, it is essential to note that no single dataset can fully encapsulate the diversity and complexity of actual network
environments. Therefore, researchers should endeavor to explore additional datasets or collect their data to further validate the performance of their AI-based IDS in various scenarios [57].

5. RESEARCH TRENDS AND FUTURE PROJECTS

The domain of AI-based Intrusion Detection Systems (IDS) in Software-Defined Networking (SDN) is continuously progressing, with research trends and future projects catering to tackling emerging challenges and augmenting the security of network environments. In this discourse, we shall lucidly deliberate on current trends and potential future projects.

1. Advanced Machine Learning Techniques: The research community is delving into advanced machine learning techniques, including deep learning, reinforcement learning, and ensemble methods, to amplify the detection capabilities of IDS in SDN. These techniques have the potential to enhance the precision and robustness of IDS models by proficiently capturing intricate patterns and behaviors in network traffic [36].

2. Explainable AI and Interpretability: The assurance of transparency and interpretability of AI-based IDS models is increasingly becoming crucial. Researchers are concentrating on developing models that provide precise detection and offer explanations or justifications for their decisions. This empowers network administrators to comprehend the reasoning behind identified threats, improving trust and aiding decision-making processes. [40].

3. Adversarial Attack Detection: Adversarial attacks aim to bypass detection by manipulating or obfuscating network traffic. Future projects focus on designing IDS models that are resilient to such attacks. Techniques such as adversarial training, anomaly detection, and adversarial sample detection are being explored to enhance the robustness of IDS models against evolving attack strategies [58].

Incorporating Contextual Information: IDS models can benefit from integrating contextual information about network behavior, user behavior, and system configurations. Researchers are studying methods to amalgamate such contextual information into detection, enhance accuracy and reduce false positives [59].
6. CONCLUSION

The goal of ongoing research, collaboration, and innovation in this field is to develop more resilient, adaptable, and efficient IDS models capable of identifying unknown and evolving threats in real-time [38]. The employment of sophisticated machine learning techniques has facilitated the refinement of the accuracy, robustness, and efficiency of IDS models. The assessment and performance analysis of AI-based IDS has furnished valuable perspectives into the efficacy of various algorithms and methodologies. Sophisticated machine-learning techniques have facilitated refining IDS models' accuracy, robustness, and efficiency [60]. These metrics serve as quantifiable measures to appraise the efficiency of the models in identifying and classifying network intrusions.

Moreover, research trends and prospective projects underscore the continual efforts to address emerging challenges and push the boundaries of IDS in SDN. These comprise advancements in explainable AI, integration of contextual data, detection of adversarial attacks, and assimilation of blockchain technology [61]. The objective is to devise IDS models that are more resilient, adaptable, and proficient in detecting unknown and evolving threats in real time. The selection of datasets utilized in research is crucial to the representativeness and dependability of the findings. Standardized datasets, including NSL-KDD, UNSW-NB15, and AWID, are widely used to validate, test, and train IDS models. However, researchers ought to continue exploring additional datasets and real-life applications to further validate the performance of their models in diverse network environments. AI-based IDS in SDN holds immense potential in enhancing network security and shielding against increasingly sophisticated threats [62]. Continuous research, cooperation, and creativity in this area will add to the evolution of IDS solutions that are more effective and intelligent, ultimately ensuring that modern network infrastructures are resilient and secure.
REFERENCES


Figure 1: A Block Diagram of IDS [15].
Figure 2: The Block Diagram of SDN Architecture [27].