Survey of Application of Machine Learning Methods in The Development of Network Intrusion Detection and Prevention Systems

Juliana Nkafu

Juliana.nkafu@gmail.com

Jun Liu

jun.liu@und.edu

School of Electrical Engineering and Computer Science
College of Engineering and Mines
University of North Dakota
Grand Forks, 58202

Abstract

Attacks targeting networks are increasing over time as Internet technology is widely adopted to foster communication in a plethora of professional and personal tasks. Attacks seek to damage and disrupt the integrity and confidentiality of connections and information exchanges. The ever-growing threats of cyber-attacks demand the urgency of developing robust security defense systems to protect business and client data. The primary goals of network defense systems are to identify, defend, and recover from network assaults. The core of network defense systems is the collective techniques for detecting and mitigating network intrusion. Network defense systems can be categorized into network intrusion detection systems (NIDS) and network intrusion prevention systems (NIPS). Network intrusion detection and prevention techniques can be categorized based on the approaches of detect network threats, the approaches of mitigating the threats, or a combination of both. The research and development on network intrusion detection and prevention techniques highly relies on the availability of representative security-related network datasets. Benchmark datasets are a good basis to evaluate and compare the quality of different network intrusion detection systems. Benchmark datasets with labeled data points of "normal" or "attach" serves as the important input to evaluate the quality of intrusion detection techniques to distinguish correctly detected attacks from false alarms. ML and DL have been used to improve IDS detection accuracy and reduce false positives. Our paper gives a survey of the application of machine learning and deep learning techniques in network intrusion detection, together with a list of available cybersecurity datasets used for model training. Network datasets labeled with information about malicious events are.

Keywords: *Network defense system, Intrusion Detection Systems, Intrusion Prevention Systems, Security-related network datasets, Machine Learning, Deep Learning*

1 Introduction

In recent year, cybercriminals launched a wave of cyberattacks that were not only highly coordinated, but also far more sophisticated than ever before. The emerging cloud evolution technologies have brought remarkable evolutions in network technology where different applications, services, and computing and storage resources are offered on demand to many users via the internet. Such an exponential growth in network technologies has offered many advantages and has improved communications. Although the Internet facilitates connection and communication, the integrity and confidentiality of these connections and information exchanges can be violated and compromised by attackers seeking to damage and disrupt network connections and network security. Each emerging network technology presents new security challenges and triggers the need for the development of detection tools and countermeasures to meet new demands.

Network attacks have become more sophisticated, and the foremost challenge is to identify unknown and obfuscated attacks as these authors use different evasion techniques for information concealing to prevent detection by an IDS. In the past, cybercriminals primarily focused on bank customers, robbing bank accounts, or stealing credit cards (Symantec, 2017). The new generation of attackers has become more ambitious and is targeting the banks themselves, sometimes trying to take millions of dollars in a single attack (Symantec, 2017). According to the purplesec.us report; on average, a malware attack costs a company over $2.5 million (including the time needed to resolve the attack). Individuals of phishing scams lost $225 on average. High profile incidents of cybercrime have demonstrated the ease with which cyber threats can spread internationally, as a simple compromise can disrupt a business' essential service or facilities. Many cybercriminals around the world are motivated to steal information, illegitimately receive revenues, and find new targets. For this reason, the detection of zero-day attacks has become the highest priority.

Several techniques for handling and classifying network traffic attacks have been proposed over the years. One approach is port-based, which involves identifying port numbers among those registered with the Internet Assign Number Authority (IANA). However, as the number of applications has grown, so has the number of unpredictable ports, and this technique has proven to be ineffective. This technique excludes account applications that do not register their ports with the IANA and use dynamic port numbers. Another technique proposed is the payload-based technique, also known as deep packet inspection (DPI), in which the contents of network packets are observed and compared to an existing set of signatures stored in a database. This method is more accurate than the port-based technique, but it does not work with network applications that use encrypted data. Behavioral classification techniques examine all network traffic received by the host to determine the type of application. The Network traffic patterns can be analyzed graphically as well as by looking at heuristic data such as transport layer protocols and the number of distinct ports contacted. Although behavioral techniques produce good results by detecting unknown threats, they are resource intensive and prone to false positives. Another technique, known as the rationale-based or statistical technique, looks at the statistical characteristics of traffic flow, such as the number of packets and the maximum, mean, and minimum packet size. Because these measurements are unique to each application, these statistical characteristics are used to identify different applications. However, there is an increasing need to combine this approach with techniques that can improve accuracy and speed up the classification of statistical patterns. Correlation-based classification groups packets into flows, or groups data packets with the same source and destination IP, port, and protocol. These are classified based on the correlation of network flows. Multiple flows are typically combined into a Bag of Flows (BoF). Although this technique outperforms statistical techniques because it eliminates feature redundancy, it has a high computational overhead for feature matching. As a result, the need to develop techniques to overcome the rising challenges persists.

The concepts of intelligent techniques, namely machine learning (ML) and deep learning (DL), became popular at the beginning of the twenty-first century. Researchers widely agreed that these techniques, which focus on using statistical methods and data to make computers think like humans, could increase the calculation potential. To address the limitations of non-intelligent techniques, computer scientists began to use intelligent techniques in network security. In network security, ML or DL algorithms can be trained on network data to distinguish between normal and malicious traffic and thus protect the network from intruders. Furthermore, if the network traffic is malicious, algorithms can be trained to identify the type of attack and take appropriate action to prevent the attack. The model can be taught to prepare individual defensive reactions by analyzing previous cyber-attacks.

This paper focuses on surveying the intelligent methods in network security can be useful in large businesses, organizations, law enforcement agencies, and banks that store sensitive information, as well as in personal networks. There are three significant contributions made by this article. (i) We did a systematic analysis to choose recent journal publications on various ML- and DL-based NIDS published during the last five years (2019-April 2022). (ii) We conducted a comprehensive examination of each publication and discussed its distinct characteristics, including its proposed methodology, evaluation criteria, network attack types, and datasets used. (iii) We did a review on the Network Intrusion and Prevention systems. (iv) On the basis of these observations, we presented recent trends in the use of AI approaches for NIDS, emphasized significant problems in ML-/DL-based NIDS, and outlined a variety of future directions in this crucial sector.

2 Network Attacks

For decades, networking technologies have been used to improve data transfer and circulation. Their continuous improvement has facilitated a wide range of new services. The utilization of mobiles is turning out to be an important component in our everyday life. An ever-increasing number of clients across the world rely upon their mobiles to trade messages, manage their personal documents, browse their emails. Additionally, mobiles facilitate online shopping which provokes clients to type their credit card numbers, security codes, usernames, and passwords. The goal of computer network defense (CND) is to prevent network intrusions that could lead to service/network denial, degradation, or disruptions by employing a variety of processes and defensive mechanisms that rely on computers and the internet. The smart city is one of the fastest growing fields. The fundamental goal of any smart city is improving the citizen's quality of life by offering a direct association to the administering body and providing better management to traffic, water, energy, air, waste, and more. Because of the variety of components in smart cities, security issues have become a significant concern.

A network attack is an approach to hurt, reveal, change, destroy, steal, or obtain illegal access to a network system resource. The attack could come from inside (internal attack) or from outside (external attack).

Existing review articles e.g., such as (Buczak & Guven, 2016; Axelsson, 2000; Ahmed et al., 2016; Lunt, 1988; Agrawal & Agrawal, 2015)) focus on intrusion detection techniques or dataset issue or type of computer attack and IDS evasion. The highly cited survey by Debar et al. (Debar et al., 2000) surveyed detection methods based on the behavior and knowledge profiles of the attacks. The major types of attacks can be categorized in the following list.

- **DDoS Attacks**

  These are attacks that attempt to disrupt the availability of service. Since distributed denial of service is easy to launch but not easy to detect, as in most cases the attacks traffic is very similar to legitimate traffic. DDoS attacks often originate from multiple sources, making them difficult to mitigate. To flood the target with traffic, the attacker uses "zombies"—compromised computers. HTTP requests, fake packets, and junk data can be this traffic. (Baek et al, 2019) conducted a study providing a model for assessing and identifying DDoS assaults on the network-level and service-levels of the bitcoin ecosystem. The dataset comprised of authentic DDoS attacks and included the impacted service, attack date, service type, number of postings, etc. The researchers collected statistical data such as maximum, minimum, total, and standard deviation from the Bitcoin block data. The researchers utilized PCA to extract features. MLP was used to identify DDoS, and the training set, validation set, and testing set were split 6:2:2 respectively.

- **Insider Threats**

  The term "insider threat" refers to a security risk posed to an organization by insiders having access to sensitive information, systems, or assets, such as employees, contractors, or business partners. This type of danger can manifest in a variety of ways, including inadvertent data breaches, malicious attacks, the theft of sensitive information, and the introduction of malware into the organization's systems. Because insider threats frequently include persons with high degrees of access and trust, it is easier for them to circumvent security measures and do damage. (Yuan et al, 2018). [12] utilized LSTM and CNN approaches to develop a model for detecting insider threats. They used the model to the CERT insider threat v4.2 dataset [13], which consisted of 32 M log lines, of which 7323 represented unusual activity. This edition of the CERT dataset contains a greater number of examples of insider threats than previous versions. The train-to-test ratio was 70% to 30%. The researchers initially extracted user behavior using LSTM, then extracted temporal characteristics and generated feature vectors. The researchers then turned the feature vectors into matrices of fixed size.

- **Phishing Attacks**

  Phishing assaults are a type of social engineering attack that seeks to acquire sensitive information, such as login credentials or financial information, by convincing victims they are talking with a reputable source, such as a bank or online service provider. These assaults are frequently delivered via email, text message, or phone call and may look to originate from a reputable entity, such as a bank or online business, requesting sensitive information or login credentials. The attacker may also include a link in the message that links to a phishing website that appears legitimate and requests sensitive information. (Mohammad et al,) built a self-structuring neural network based on ANN to recognize phishing website attacks. Phishing-related traits are essential for detecting highly dynamic web sites; hence, the network's architecture must be continuously enhanced. The suggested method solves this issue by automating the process of network architecture and displaying a high tolerance for noisy input, fault tolerance, and significant prediction accuracy. This was accomplished by accelerating the learning rate and adding neurons to the hidden layer. The objective of the constructed model was to achieve generalization ability, which necessitates that the classification accuracy throughout training and testing be as comparable as possible.

- **Malware**

  Malware is software that harms or exploits a computer system or network. Viruses, worms, trojan horses, ransomware, spyware, and adware are malware. Viruses spread by attaching themselves to emails or other files. Unlike viruses, worms replicate without attaching to files. Trojan horses are malware that masquerade as harmless software and fool users into downloading and installing them. Ransomware encrypts files and demands payment to decrypt them. Spyware steals passwords and login credentials from victims' computers. Adware shows unwelcome ads on victims' computers. Using the RF technique and the Kyoto 2006+ (Song et al, 2011) dataset, (Park et al, 2018) examined the recognition performance of several forms of attacks, including IDS, malware, and shellcode (total size 19.8 GB). The dataset contained three types of class: attack, shellcode, and normal.

- **Zero-Day Attacks**

  A zero-day attacks is a sort of cyberattack that exploits a previously unknown software application or operating system vulnerability. The phrase "zero-day" refers to the fact that the vulnerability has not been identified or revealed to the public; hence, the producer of the program has had zero days to patch the weakness and prevent it from being exploited. Zero-day attacks are especially perilous because they exploit unpatched vulnerabilities, allowing attackers to infiltrate a target's systems and steal sensitive data or install malware. These attacks can be launched by nation-states, criminal groups, or individual hackers for several goals, including cyber espionage, sensitive data theft, and financial gain. Several researchers have, surprisingly, focused on discovering zero-day attacks. (Beaver et al, 2013) conducted one such investigation using machine learning techniques that can distinguish between normal and malicious communications.

3 Network Intrusion Detection and Prevention Systems

Network security has recently received an enormous attention due to the mounting security concerns in today’s networks. computer security has become essential as the use of information technology has become part of our daily lives. Undoubtedly, IoT devices are vulnerable to various security attacks. There is a serious requirement for IDSs to secure IoT gadgets against security vulnerabilities.

3.1 Intrusion Detection System (IDS)

An IDS intensely monitors malicious network activities and notifies officials if an attack is detected with no prevention abilities. Signature-based and anomaly-based detection are the two most prevalent approaches used by IDS to identify threats. Typically, IDS works in three steps: monitoring, detecting, and warning. Firstly, it monitors the network traffic or the system. Secondly, it analyzes and identifies the pattern of connections and intrusion behaviors accordingly to the characteristics of the used algorithm. Finally, it generates an alarm immediately when detecting a suspicious activity for investigation. On the other hand, anomaly-based procedures attempt to differentiate malicious traffic from real traffic based on a change in the network traffic; thus, they can detect unknown threats. On the other hand, anomaly-based procedures attempt to differentiate malicious traffic from real traffic based on a change in the network traffic; thus, they can detect unknown threats.

- **Signature-based intrusion detection systems (SIDS)**
  Signature intrusion detection systems (SIDS) are based on pattern matching techniques to find a known attack; these are also known as Knowledge-based Detection or Misuse Detection. matching methods are used to find a previous intrusion. In other words, when an intrusion signature matches with the signature of a previous intrusion that already exists in the signature database, an alarm signal is triggered. SIDS usually gives an excellent detection accuracy for previously known intrusions. However, SIDS has difficulty in detecting zero-day attacks because no matching signature exists in the database until the signature of the new attack is extracted and stored. As a result of having to establish a new signature for each alteration, the efficiency of signature-based systems is drastically reduced. Moreso, the increasing rate of zero-day attacks has rendered SIDS techniques progressively less effective because no prior signature exists for any such attacks (Symantec, 2017).

- **Anomaly-based intrusion detection system (AIDS)**
  In AIDS, a normal model of the behavior of a computer system is created using machine learning, statistical-based or knowledge-based methods. Any significant deviation between the observed behavior and the model is regarded as an anomaly, which can be interpreted as an intrusion. The classification is based on heuristics or rules, rather than patterns or signatures. This category of strategies assumes harmful activity is different from user behavior; intrusions are anomalous user behavior. AIDS development involves training and testing. The training phase uses the typical traffic profile to create a model of normal behavior, and the testing phase uses a new data set to assess the system's ability to generalize new intrusions. AIDS can detect zero-day attacks without a signature database by analyzing abnormal user behavior (Alazab et al., 2012). Hence, AIDS has advantages. They can initially detect organizational malfeasances; An alarm is triggered whenever an intruder makes suspicious transactions in a compromised account. Second, because the system uses individualized profiles, cybercriminals can't detect routine behavior without triggering an alert.

- **Host-based IDS (HIDS)**
  A host-based intrusion detection system (HIDS) is a security solution that monitors and analyzes activity on individual computer systems or hosts in search of indicators of malicious behavior or unauthorized access. Unlike network-based intrusion detection systems (NIDS), which monitor

network traffic, host-based intrusion detection systems (HIDS) operate at the host level and are capable of detecting both internal and external threats. Comparing the present state of a host to its baseline or expected state and searching for deviations or anomalies that may signal an intrusion or breach is how HIDS software normally operates. Unauthorized changes to system files, attempts to access privileged resources or data, and the installation or execution of malicious software are examples of activities that may generate a HIDS warning. HIDS can detect insider attacks that do not involve network traffic. HIDS can provide an additional layer of defense against a wide variety of cyber threats, hence enhancing the security posture of individual hosts or systems.

- **Network-based IDS (NIDS)**
  A network-based intrusion detection system (NIDS) is a security solution that analyzes network data for indications of malicious or suspicious behavior. Unlike host-based intrusion detection systems (HIDS), which focus on specific hosts or systems, NIDS functions at the network level, examining network traffic in search of patterns or behaviors that may indicate a security issue. NIDS typically operate by recording network data in real-time and evaluating it for indicators of malicious activity or policy violations. This may involve recognizing known attack signatures, examining traffic patterns for anomalies, or employing machine learning techniques to discover patterns that may signal an attack. NIDS can be deployed as a standalone device or as a component of a larger network security architecture. Certain NIDS solutions are intended for integration with other security technologies, such as firewalls and intrusion prevention systems, to provide a more comprehensive protection against cyber-attacks.

3.2 Intrusion Prevention System (IPS)

The Intrusion Prevention System, often known as intrusion detection and prevention systems, is abbreviated as "IPS" (IDPS). It does a continual search across the network to identify any unauthorized or rogue control points that may be present. These points are identified based on changes in behavior. The system will automatically take preventative actions to deal with the dangers and protect itself from further damage. The protection of a network against harmful or unwanted packets and assaults is the primary purpose of an intrusion detection and prevention system (IDPS). IDPS is more effective than IDS because in addition to detecting risks, it is also able to respond appropriately to such threats. There are two different kinds of intrusion detection and prevention systems (IDPS): network-based intrusion detection and prevention systems (NIDPS), which examine the network protocol Sensors 2021, 21, 7070 6 of 43 to identify any suspicious activities; and host-based intrusion detection and prevention systems (HIDPS), which are utilized to monitor host activities for any suspicious events that may occur within the host. ML or DL based intelligent techniques ML and DL Can be adopted for effectively and efficiently identify network attacks.

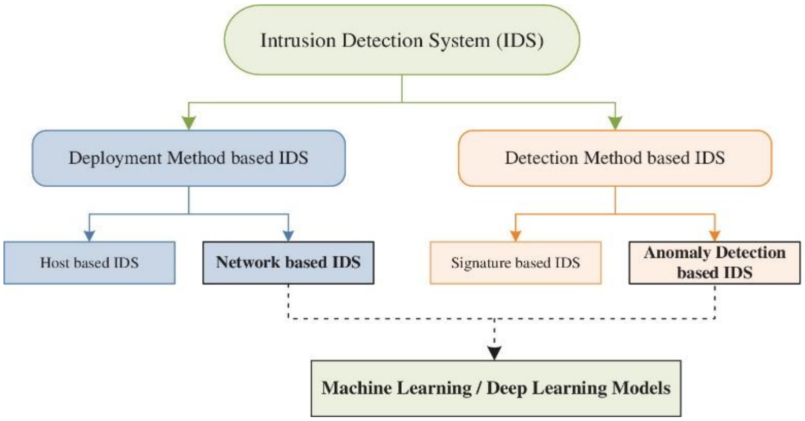

Intrusion Detection System (IDS)
Deployment Method based IDS
Detection Method based IDS
Host based IDS
Network based IDS
Signature based IDS
Anomaly Detection based IDS
Machine Learning / Deep Learning Models

Figure 1: Network Intrusion Detection Systems Classification Taxonomy.

4 Applying ML Techniques in the Design of NDS and IDS

People are always looking for methods, tools, or techniques that reduce the amount of effort required to perform a task efficiently. In Machine Learning, algorithms are programmed to attempt to self-learn based on their past experiences. After gaining knowledge from previous experiences, algorithms become quite capable of reacting and responding to conditions for which they were not explicitly programmed. Consequently, Machine Learning contributes significantly to intrusion detection. It attempts to recognize previously unrecognized hidden patterns that aid in intrusion detection. Machine Learning approaches for intrusion detection are so widespread and popular.

- **Improved Accuracy**
  The algorithms that comprise machine learning can examine huge volumes of data, recognize patterns, and learn from these observations. This can lead to more accurate intrusion detection, with fewer false positives and false negatives because of the process. It quickly examines and processes data to extract new patterns from it. For humans to analyze the data will require a substantial amount of time, which will increase proportionally to the quantity of data. Rule-based intrusion detection systems rely on established rules to determine which types of actions are regarded safe and which should raise a red flag. Rule-based method is inefficient due to the time-consuming nature of writing these rules for various instances. Machine Learning-based Intrusion Detection algorithms succeed in learning from existing patterns and can automatically recognize new patterns. And it accomplishes all of this in a fraction of the time required by rule-based systems.

- **Adaptability**
  Algorithms that are learned through machine learning can adjust to new kinds of threats and changing settings. The ability of machine learning models to identify new threats can progress in tandem with the development of novel intrusion techniques by cybercriminals. In addition, machine learning models are continually updated and retrained to accommodate the emergence of new types of assaults and the evolution of threat landscapes.

- **Scalability**
  Because machine learning models can process massive amounts of data, these models are scalable and therefore suitable for use in enterprise-level intrusion detection systems. This may be of particular benefit to companies that have a significant number of endpoints and network devices. The goal of intrusion detection in network security is to identify and respond to potential security breaches in real-time. Their ability to quickly and efficiently process and analyze large amounts of data makes them well-suited to handling the high volume of network traffic that is typical in modern networks.

4.1 Intrusion Data Sources

Intrusion data sources can be categorized by the methodologies used to detect intrusions (signature-based or anomaly-based) or by the input data sources used to detect anomalous behaviors (host-based or network-based). These methods use software, hardware, or both.

Datasets can test new methods. The dataset's amount and quality affect an IDS's performance. Packet-based, flow-based, or other formats can collect network traffic. Common features as evaluation bases assist researchers find suitable data sets for their evaluation scenarios. Network and transport protocols evaluate packet-based data. TCP, UDP, ICMP, and IP are the main protocols. Pcaps of these protocols include the payload. TCP, a reliable transport protocol, uses metadata including sequence numbers, acknowledgement numbers, TCP flags, and checksum values. Flow-based data have no payload and are created by collecting all packets that arrive within a defined time frame that share specific properties into a single flow. It contains the first view date, duration, and transport protocols and is used to match flow-based data attributes. Flows may be unidirectional or bidirectional. Netflow, IP-FIX, sFlow, and OpenFlow are flow-based formats (Mckeown et al, 2008). Other data includes packet- and flow-free data collections. This category includes flow-based data sets supplemented with packet-based or host-based log files.

Table 1: Overview of cybersecurity datasets

Dataset Name	Year	Attack types	Observation
DARPA	1998/99	Dos, R2L, U2R, Probe	Include irregular distribution of attack data instances Do not represent real network traffic
KDD Cup 99	1999	Dos, R2L, U2R, Probe	Suffer from redundant records and duplicate data samples.
Kyoto 2006+	2006/2008	DoS, Probe, U2R, R2L	While the Kyoto 2006+ dataset includes a diverse range of attack types, there may be other types of attacks that are not represented in the dataset, such as advanced persistent threats (APTs) or insider attacks.
NSL-KDD	2009	Dos, R2L, U2R, Probe	Lack of redundant records Limited number of attack types

UNSW-NB15	2015	Backdoors, portscans, DoS, Exploits, Spam, Reconnaissance, fuzzers, generic, Shellcode,Worms	Have list of new attacks and updated continuously.
CIC-IDS2017	2017	Brute force, Dos, DDoS, Portscan, Web, Botnet, Infiltration	contain some redundant data records Include attacks that resembles the real-world data
CSE-CIC-IDS 2018	2018	Brute force, Dos, DDoS, Portscan, Web, Botnet, Infiltration	Generate the dataset with the help of network profiles List a new scope of attacks produced from real network traffic
DARPA AI Next	2019	Malware infections, C2, Data exfiltration	The dataset was created with the specific goal of testing the performance of AI systems, and as such, it may not include all types of cyber threats or operations.
IoT-23	2020	Benign and malicious traffic	Have list of new attacks and updated continuously.

4.2 Challenges of Applying Machine Learning methods in Designing IDS

Since the accuracy of ML methods depends on the quality of data, it is crucial to provide appropriate datasets for training and testing phases. There's no doubt that ML methods have significantly improved the IDS landscape by rapidly identifying and frustrating attacks. However, Because of the continuous growing sophistication of the threats, ML remains incapable to keep up with this flow of threats due to some reasons outlined below

- **Lack of sufficient data**
  It is necessary to offer suitable datasets for the training and testing phases, as the precision of machine learning techniques is directly proportional to the quality of the data. Due to a lack of relevant datasets, it is impossible to conduct a security threat assessment or develop an efficient defense plan. Older training datasets are a significant challenge for any approach because they result in a detection performance that is only moderate when applied to fresh dangers. In addition to this, one of the challenges that machine learning specialists need to overcome is an imbalanced dataset. When a dataset has an uneven distribution of classes, for example when the ratio of harmful to benign samples is 2 to 40, we say that the dataset is imbalanced.

- **Labeled Sample Shortages**
  Because of their low cost and the ease with which they can be trained and put into practice; supervised learning methods are the most common type of machine learning approach utilized for intrusion detection systems (IDS). Yet, the most significant drawback of this machine learning category is that it cannot be trained without labeled examples. Unfortunately, there are not a lot of datasets that have labels, and the manual development of these labels is a time-consuming and expensive operation. It is widespread practice to use external whitelists and blacklists when labeling products, although the accuracy of these lists cannot be guaranteed, and they may also be of inferior quality.

- **Approaches of Attacks**
  Due to the strategic nature of the attacks and the fact that they are always adjusting their techniques, the application of machine learning is made more difficult. The defender makes repeated efforts to guard his holdings using all the resources at his disposal. So, in most situations, he sets up numerous layers of defensive systems and waits for some indication of an attack before acting. While it is true that most of the time, the attacker is aware of his objective as well as the type of defense that needs to be breached. He also has the advantage of knowing the exact time as well as the measures that will be taken during the assault. In addition, those who carry out attacks are continually developing new methods and making use of the most recent technological advancements. This includes both artificial intelligence and machine learning. To emerge victorious from the conflict, the defender needs to go above and beyond the simple duty of warding off attacks and instead focus on rendering them impossible

4.3 Different ML Techniques Applied in the Design of IDS

Machine learning is the process of extracting knowledge from large quantities of data. Machine learning models comprise of a set of rules, methods, or complex "transfer functions" that can be applied to find interesting data patterns, or to recognize or predict behavior (Dua & Du, 2016). This learning could either be supervised, unsupervised, semi-supervised, ensembled or hybrid. The goal of using machine learning techniques is to create IDS with improved accuracy and less requirement for human knowledge, which are some of the reasons why machine learning is popular these days.

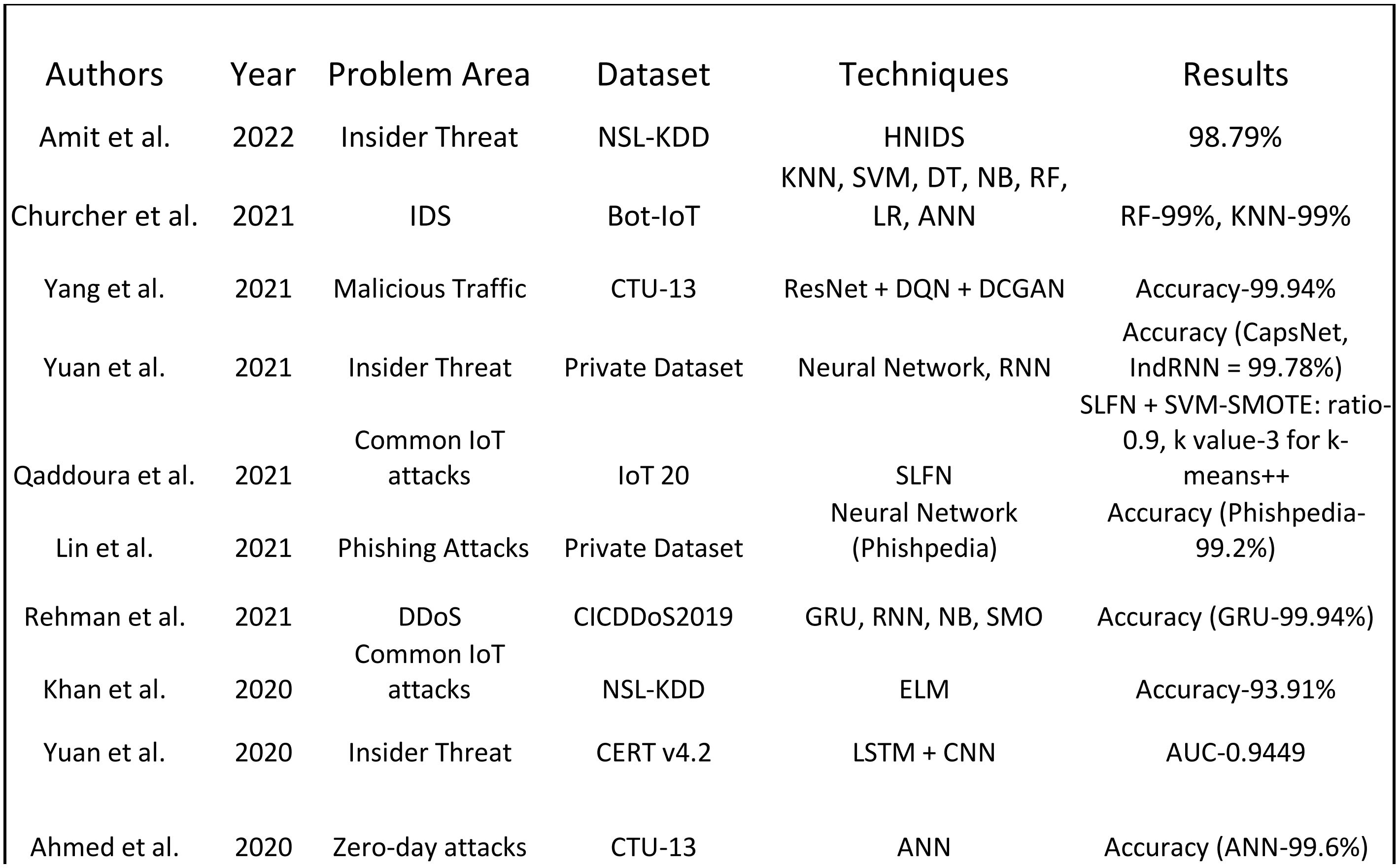
Supervised learning-based IDS techniques detect intrusions by using labeled training data. This approach usually consists of two stages, namely training and testing. In the training stage, relevant features and classes are identified and then the algorithm learns from these data samples. Each record is a pair, containing a network or host data source and an associated output value (i.e., label), namely intrusion or normal. Next, feature selection can be applied for eliminating unnecessary features. Using the training data for selected features, a supervised learning technique is then used to train a classifier to learn the inherent relationship that exists between the input data and the labelled output value. In the testing stage, the trained model is used to classify the unknown data into intrusion or normal class. The resultant classifier then becomes a model which, given a set of feature values, predicts the class to which the input data might belong. The performance of a classifier in its ability to predict the correct class is measured in terms of several metrics.
There are many classification methods such as decision trees, rule-based systems, neural networks, support vector machines, naïve Bayes, and nearest-neighbor. Each technique uses a learning method to build a classification model.

Unsupervised learning is a form of machine learning technique used to obtain interesting information from input datasets without class labels. The input data points are normally treated as a set of random variables. A joint density model is then created for the data set. In supervised learning, the output labels are given and used to train the machine to get the required results for an unseen data point, while in unsupervised learning, no labels are given, and instead the data is grouped automatically into various classes through the learning process. In the context of developing an IDS, unsupervised learning means, use of a mechanism to identify intrusions by using unlabeled data to a train the model.

Semi-supervised learning falls between supervised learning (with totally labelled training data) and unsupervised learning (without any categorized training data). Researchers have shown that semi-supervised learning could be used in conjunction with a small amount of labelled data classifier’s performance for the IDSs with less time and costs needed. This is valuable as for many IDS issues, labelled data can be rare or occasional (Ashfaq et al., 2017). Several different techniques for semi-supervised learning have been proposed, such as the Expectation Maximization (EM) based algorithms (Goldstein, 2012), self-training (Blount et al., 2011; Lyngdoh et al., 2018), co-training (Rath et al., 2017), Semi-Supervised SVM (Ashfaq et al., 2017).

Combining machine learning techniques improves predicted performance. Boosting, Bagging, and Stacking are ensemble approaches. Conventional IDSs cannot be changed, cannot recognize new malicious threats, have low accuracy, and high false alarms. AIDS's false-positive rate. SIDS and AIDS form hybrid IDS. Hybrid IDS overcomes SIDS and AIDS. (Farid et al, 2010) suggested a hybrid IDS employing Naive Bayes and decision trees to detect 99.63% of KDD'99 datasets.

Table 2. Summaries of reviewed papers.



Authors	Year	Problem Area	Dataset	Techniques	Results
Amit et al.	2022	Insider Threat	NSL-KDD	HNIDS	98.79%
Churcher et al.	2021	IDS	Bot-IoT	KNN, SVM, DT, NB, RF, LR, ANN	RF-99%, KNN-99%
Yang et al.	2021	Malicious Traffic	CTU-13	ResNet + DQN + DCGAN	Accuracy-99.94%
Yuan et al.	2021	Insider Threat	Private Dataset	Neural Network, RNN	Accuracy (CapsNet, IndRNN = 99.78%)
Qaddoura et al.	2021	Common IoT attacks	IoT 20	SLFN	SLFN + SVM-SMOTE: ratio-0.9, k value-3 for k-means++
Lin et al.	2021	Phishing Attacks	Private Dataset	Neural Network (Phishpedia)	Accuracy (Phishpedia-99.2%)
Rehman et al.	2021	DDoS	CICDDoS2019	GRU, RNN, NB, SMO	Accuracy (GRU-99.94%)
Khan et al.	2020	Common IoT attacks	NSL-KDD	ELM	Accuracy-93.91%
Yuan et al.	2020	Insider Threat	CERT v4.2	LSTM + CNN	AUC-0.9449
Ahmed et al.	2020	Zero-day attacks	CTU-13	ANN	Accuracy (ANN-99.6%)

Letteri et al.	2020	Malware Attack	MTA KDD 19	MLP using AE optimization or RRw optimization	Accuracy (MLP with RRw opt.-99.60%)
Kim et al.	2020	DDoS	KDD-99, CICIDS2018	CNN, RNN	Accuracy (CNN-99% or more)
Alrashdi et al.	2019	Common IoT attacks	UNSW-NB15	RF	Accuracy (ML-99.34%)
Zhang et al.	2019	IDS	NSL-KDD	AE	F-Score-76.47% Recall-79.47%
Hu et al	2019	Insider Threat	Private Dataset	CNN	FAR-2.94% FRR-2.28%
Pektas et al.	2019	Botnet Attacks	ISOT HTTP, CTU-13	MLP + LSTM	ISOT: F score-98.8% CTU: F score-99.1%
Nguyen et al.	2018	IDS	UNSW-NB15, KDD-99, NSL-KDD	NNET	Accuracy (KDD-99-97.11%)

5. **Conclusion**

Network security is a major concern for individuals, businesses, and governments. With the current digital explosion, network security is essential to secure society's acceptance of the tens of thousands of services that rely on the network, the backbone of digital life. Hence, network security is essential. This study reviews IDS machine learning classification techniques. Researchers have used SVM, Nave Bayes, Neural Network, Gradient Boosted Tree, Decision Tree, k-nearest neighbors, multinomial randomness, forest classifier, stochastic gradient descent, and ensemble classifiers. SVM, Random Forest, and CNN can identify with high accuracy. We examined the most popular public datasets for IDS research, their data gathering methods, evaluation findings, and constraints. Newer and more complete malware activity datasets are needed because normal activities vary frequently and may lose effectiveness over time. DARPA/KDD99 does not include new malware operations. As these 1999 datasets are publicly available and no other appropriate datasets exist, testing is limited to them. These benchmarks no longer represent modern zero-day attacks. Combination techniques will be tested on the same dataset to improve detection and reduce false positives.

References

[Symantec, 2017] Symantec, "Internet security threat report 2017," April, 7017 2017, vol. 22.

[Goli et al., 2018] Y. D. Goli, R. Ambika, Network Traffic Classification Techniques-A Review. In Proceedings of the International Conference on Computational Techniques, Electronics and Mechanical Systems, CTEMS 2018, Belgaum, India, 21–22 December 2018; pp. 219–222.

[Buczak et al., 2016] A. Buczak, E.Guven, (2016) A survey of data mining and machine learning methods for cyber security intrusion detection. IEEE Communications Surveys & Tutorials 18(2):1153–1176.

[Park et al, 2018] K. Park, Y. Song, Y. G. Cheong, Classification of attack types for intrusion detection systems using a machine learning algorithm. In Proceedings of the 2018 IEEE Fourth International Conference on Big Data Computing Service and Applications (BigDataService), Bamberg, Germany, 26–29 March 2018.

[Song et al, 2011] J. Song, H. Takakura, Y. Okabe, M. Eto, Inoue, K. Nakao, Statistical analysis of honeypot data and building of Kyoto 2006+ dataset for NIDS evaluation. In Proceedings of the 1st Workshop on Building Analysis Datasets and Gathering Experience Returns for Security, BADGERS 2011, Salzburg, Austria, 10 April 2011.

[Beaver et al, 2013] J. M. Beaver, C. T. Siymons, R. E. Gillen, A learning system for discriminating variants of malicious network traffic. In Proceedings of the Eighth Annual Cyber Security and Information Intelligence Research Workshop, Oak Ridge, TN, USA, 8–10 January 2013.

[Baek et al, 2019] J. U. Baek, S. H. Ji, J. T. Park, M. S. Kim, DDoS Attack Detection on Bitcoin Ecosystem using Deep-Learning. In Proceedings of the 2019 20th Asia-Pacific Network Operations and Management Symposium (APNOMS), Matsue, Japan, 18–20 September 2019.

[Mohammad et al, 2014] R. M. Mohammad, F. Thabtah, L. McCluskey, Predicting phishing websites based on self-structuring neural network. Neural Comput. Appl. 2014, 25, 443–458.

[Yuan et al, 2018] F. Yuan, Y. Cao, Y. Shang, Y. Liu, J. Tan, B. Fang, Threat Detection with Deep Neural Network. In Computational Science—ICCS 2018; Springer: Cham, Switzerland, 2018.

[Debar et al, 2000] Dacier, M Dacier Deber, and A. Wespi, "A revised taxonomy for intrusiondetection systems," in Annales des télécommunications, 2000, vol. 55, no. 7–8, pp. 361–378: Springer.

[Alazab et al, 2012] Hobbs M. Alazab, J. Abawajy, and M. Alazab, "Using feature selection for intrusion detection system," in 2012 international symposium on communications and information technologies (ISCIT), 2012, pp. 296–301.

[Claise, 2012] Cisco Systems NetFlow Services Export Version 9. Internet Engineering Task Force 2004. doi:10.17487/RFC3954.

[Claise, 2012] Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of IP Traffic Flow Information. Internet Engineering Task Force 2008.doi:10.17487/RFC5101.

[McKeown et al, 2016] Anderson N, McKeown, T. Balakrishnan H.Parulkar G. Peterson L. Rexford J. Shenker S. Turner J, OpenFlow: enabling innovation in campus networks. ACM SIGCOMM Comput Commun Rev 2008;38(2):69–74. doi:10.1145/1355734.1355746.

Dua and X. Du, Data mining and machine learning in cybersecurity. CRC press, 2016

[Ashfaq et al, 2017] R. Ashfaq, X-Z. Wang, JZ. Huang, H. Abbas, Y-L. (2017) Fuzziness based semisupervised learning approach for intrusion detection system. Inf Sci 378:484–497.

[Goldstein, 2012] Goldstein, "FastLOF: an expectation-maximization based local outlier detection algorithm," in Pattern recognition (ICPR), 2012 21st international conference on, 2012, pp. 2282–2285: IEEE.

[Lyngdoh et al, 2018] M.Lyndoh, I. Hussain, S. Majaw, and H. K. Kalita, "An intrusion detection method using artificial immune system approach," in international conference on advanced informatics for computing research, 2018, pp. 379–387: Springer.

[Rath et al, 2017] PS. Rath, NK. Barpanda, R. Singh, S. Panda (2017) A prototype Multiview approach for reduction of false alarm rate in network intrusion detection system. Int J Comput Netw Commun Secur 5(3):49.

[Farid et al, 2010] Harbi N.Farid, M. Z. Rahman, "Combining naive bayes and decision tree for adaptive intrusion detection," arXiv preprint arXiv:1005.4496, 2010.

[Amit et al, 2022] Kumar B. Amit, Sachin Ahunja, Kumar L. Umesh, Sanjeev K. Sharma, Poongodi Manoharan., Abeer D. Algarni, Hela Elmannai, Kaamran Raahemifar, A hybrid intrusion detection model using EGA-PSO and improved random forest method. Sensor 2022, 22(16), 5986.

[Churcher et al, 2021] A Churcher, R. Ullah, J. Ahmad, Ur S. Rehman, F Masood, M. Gogate, F. Alqahtani, B. Nour, W.J. Buchanan, an experimental analysis of attack classification using machine learning in IoT networks. Sensors 2021, 21, 446.

[Yang et al, 2021] J. Yang, G. Liang, B. Li, G. Wen, T. Gao, A deep-learning- and reinforcement-learning-based system for encrypted network malicious traffic detection. Electron. Lett. 2021, 57, 363–365.

[Yuan et al, 2021] J. Yuan, G. Chen, S. Tian, X. Pei, Malicious URL detection based on a parallel neural joint model. IEEE Access 2021, 9, 9464–9472.

[Qaddoura et al, 2021] R. Qaddoura, A.M. Al-Zoubi, I. Almomani, H. Faris, A multi-stage classification approach for iot intrusion detection based on clustering with oversampling. Appl. Sci. 2021, 11, 3022.

[Qaddoura et al, 2021] R. Qaddoura, A.M. Al-Zoubi, H. Faris, I. Almomani, A multi-layer classification approach for intrusion detection in iot networks based on deep learning. Sensors 2021, 21, 2987.

[Lin et al, 2021] Y. Lin, R. Liu, M. Divakaran, J. Y. Ng, Q.Z Chan, Y. Lu, Y. Si, F. Zhang, J.S. D. Phishpedia, A Hybrid Deep Learning Based Approach to Visually Identify Phishing Webpages. In Proceedings of the 30th {USENIX} Security Symposium ({USENIX} Security 21, Online, 11–13 August 2021.

[Rehman et al, 2021] S. Rehman, M. Khaliq, S.I. Imtiaz, A. Rasool, M. Shafiq, A.R. Javed, Z. Jalil, A.K. B. Diddos, An approach for detection and identification of Distributed Denial of Service (DDoS) cyberattacks using Gated Recurrent Units (GRU). Futur. Gener. Comput. Syst. 2021, 118, 453–466.

[Yang et al, 2020] C.T. Yang, J.C. Liu, E. Kristiani, M.L. Liu, I. You, G. Pau, NetFlow Monitoring and Cyberattack Detection Using Deep Learning with Ceph. IEEE Access 2020, 8, 7842–7850.

[Zhang et al, 2019] C. Zhang, F. Ruan, L. Yin, X. Chen, L. Zhai, F.A. Liu, A Deep Learning Approach for Network Intrusion Detection Based on NSL-KDD Dataset. In Proceedings of 2019 IEEE 13th International Conference on Anti-counterfeiting, Security, and Identification (ASID), Xiamen, China, 25–27 October 2019; pp. 41–45.

[Letteri et al, 2020] I. Letteri, G. Penna, L. D. Vita, M.T Grifa, MTA-KDD'19: A Dataset for Malware Traffic Detection. 2020.

[Kim et al, 2020] Kim J. Kim, H. Kim, M. Shim, E. Choi, CNN-Based Network Intrusion Detection against Denial-of-Service Attacks. Electronics 2020, 9, 916.

[Alrashdi et al, 2019] I. Alrashdi, A. Alqazzaz, E. Aloufi, R. Alharthi, M. Zohdy, H. Ming, AD-IoT: Anomaly detection of IoT cyberattacks in smart city using machine learning. In Proceedings of the 2019 IEEE 9th Annual Computing and Communication Workshop and Conference, CCWC 2019, Las Vegas, NV, USA, 7–9 January 2019.

27. Zhang, C.; Ruan, F.; Yin, L.; Chen, X.; Zhai, L.; Liu, F. A Deep Learning Approach for Network Intrusion Detection Based on NSL-KDD Dataset. In Proceedings of 2019 IEEE 13th International Conference on Anti-counterfeiting, Security, and Identification (ASID), Xiamen, China, 25–27 October 2019; pp. 41–45.

[Hu et al, 2019] T. Hu, W. Niu, X. Zhang, X. Liu, J. Lu, Y. Liu, An Insider Threat Detection Approach Based on Mouse Dynamics and Deep Learning. Secur. Comm. Netw. 2019, 2019, 12.

[Pekta et al, 2019] A. S. Pekta, T. Acarman, Deep learning to detect botnet via network flow summaries. Neural Comput. Appl. 2019, 31, 8021–8033.

[Creech et al, 2014] G. Creech, J. Hu (2014a) A semantic approach to host-based intrusion detection systems using Contiguousand Discontiguous system call patterns. IEEE Trans Comput 63(4):807–819

[NSL-KDD, 2018] University of New Brunswick. NSL-KDD Data Set for Network-Based Intrusion Detection Systems. NSL-KDD Dataset. 2018.

[Song et al, 2006] J. Song, H. Takakura, Y. Okabe, M. Eto, D. Inoue, K. Nakao, Statistical analysis of honeypot data and building of Kyoto 2006+ dataset for NIDS evaluation. In Proceedings of the 1st Workshop on Building Analysis Datasets and Gathering Experience Returns for Security, BADGERS 2011, Salzburg, Austria, 10 April 2011.

[Nguyen et al, 2018] K.K. Nguyen, D.T. Hoang, D. Niyato, P. Wang, D. Nguyen, E. DutkiewicCyberattack detection in mobile cloud computing: A deep learning approach. IEEE Wirel. Commun. Netw.Conf. WCNC 2018, 2018, 8376973.