A Survey on Wireless Network Intrusion Detection

S. Maheswari[1], J. C. Miraclin Joyce Pamila[2]

[1]PG Scholar , Dept. of CSE, Government College of Technology, Coimbatore, India
[2]Professor, Dept. of CSE, Government College of Technology, Coimbatore, India

E-mail: [1]mahe.66976@gct.ac.in, [2]miraclin@gct.ac.in

Abstract

Artificial Intelligence (AI) discoveries have intensified in recent years as a result of the industry's widespread adoption of this technology. The important field of AI is neural networks, that allow commercial usage of capabilities that were previously unattainable through computer use. One of the domains in which neural network is widely studied for increasing general security and data privacy is IDS. Using various machine learning approaches, this article provides a complete review of recent research on neural network topologies and types of intrusion detection systems.

Keywords: Network security, CNN, Network attacks, Intrusion detection system (IDS), NSL, KDD

1. Introduction

The Internet and communication technologies have advanced at such a rapid pace that they now pervade nearly every aspect of our lives. As a result, the amount of data collected and processed has skyrocketed, ushering in the age of "big data." Since then, protecting this data and its connections has been a difficult task. Individuals and businesses may suffer serious problems if there is any fraud or sense of insecurity during data transmission. Furthermore, the complexity of the data network and the variety in attack tactics have complicated the task. As a result, experts are investigating all possible tactics and methods for maintaining a continuous connection. An IDS is an example.

Intrusion Detection System's goal is to identify malicious traffic. There are numerous techniques to implementing IDS. Anomaly detection is the most prominent of these. It is based on the detection of anomalies in traffic. Depending on the criteria used to measure traffic profile deviation, many implementations of this technique have been offered. Users'

DOI: https://doi.org/10.36548/jsws.2022.1.001
Received: 28.03.2022, received in revised form: 27.04.2022, accepted: 11.05.2022, published: 21.05.2022

basic security requirements may be met by the common wireless network authentication mechanism and firewall technology, but they have very inadequate protective capabilities when malicious attacks by professional hackers occur. Misuse detection and aberrant detection are two popular intrusion detection methods with problems include poor feature extraction, low fault detection, and a false detection rate. One of IDSys hotspots is AI- based detection techniques research as a result of their implementation in IDSys.

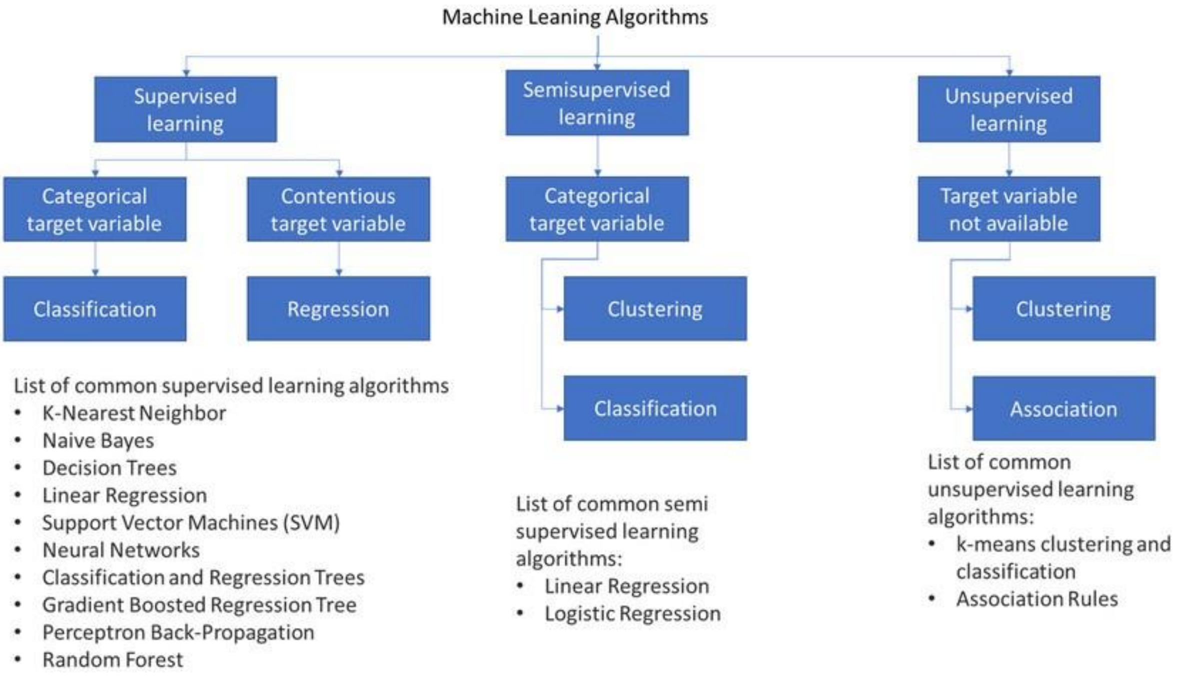
Machine Leaning Algorithms
Supervised learning
Semisupervised learning
Unsupervised learning
Categorical target variable
Contentious target variable
Categorical target variable
Target variable not available
Classification
Regression
Clustering
Classification
Clustering
Association
List of common supervised learning algorithms
• K-Nearest Neighbor
• Naive Bayes
• Decision Trees
• Linear Regression
• Support Vector Machines (SVM)
• Neural Networks
• Classification and Regression Trees
• Gradient Boosted Regression Tree
• Perceptron Back-Propagation
• Random Forest
List of common semi supervised learning algorithms:
• Linear Regression
• Logistic Regression
List of common unsupervised learning algorithms:
• k-means clustering and classification
• Association Rules

Figure 1. Machine learning algorithms

Network Based Intrusion Detection Systems (NBIDS) are the most common type of IDS (NIDS). If it senses an attack, this software network traffic analysis and warns the system administrator. HBIDS (Host Based Intrusion Detection Systems) is the second category (HIDS). The HIDS analyzes each host item separately (rather than the existing system) and alerts the host if any unusual packets are discovered. This research looks at NIDS, which is divided into two types: misuse detection and anomaly - based. To detect them, a collection of attack signatures must be well before into the abuse detection system. As a result, it is unable to identify unknown threats. From the other side, anomaly detection is based on normal usage patterns and can detect unknown dangers. Due to the process of recognizing various common use patterns, the outlier detection system has a high prone to false alarms. To put it another way, using a self-learning technique like Deep Learning (DL) models would improve the ability of the anomaly detection system to identify typical user behavior. This should assist reduce false alarms.

All machine learning algorithms are divided into shallow and deep learning in Figure1. CNN, like other types of neural networks, is a supervised deep-learning technique. R. Upadhyay and D. Pantiukhin employed it for the first time in intrusion detection in 2017. From that article through July 2021, our survey covers 81 papers that use CNN for IDS, either alone or in combination with another shallow or DL technique. No previous survey, to our knowledge, has particularly addressed CNN among the many DL techniques.

1.1 Types of Attacks

The IDS, a significant accomplishment breakthrough in the field of information security, is capable of detecting attacks, whether they are ongoing or have already occurred. The acquisition of an intervention is frequently the equivalent of a separation extractant, such as a bitwise or cross problem, i.e., determining whether network traffic behavior is acceptable or complex, or a 5 phase problem, i.e., determining whether it is common or any of the 4 types of attacks: Service Denial (DOS), Probe (Probing), and Root to Local (R2L), User To Root (U2R). The four types of attacks found in the NSL-KDD data set are as follows:

1. DoS: Syn floodings are a type of DoS attack in which the victim's resources are depleted, preventing it from processing valid requests.
2. PROBE: A probe is a type of attack in which surveillance and other investigative techniques are used to gather information on a remote victim, such as a hole scan.
3. U2R: Unauthorized access to local user privileges (root) is a type of attack in which the attacker logs into victim's system using a regular account and attempts to get root/administrator rights by exploiting the victim, such as a buffer overflow attack.
4. R2L: Remote access to a local machine, in which the attacker logs in to the remote console and acquires local access to the victim's machine.

In short, one of the main reasons for the acquisition of intervention is to improve the accuracy of class dividers in correctly spotting poor behavior. Machine learning techniques for detecting various types of attacks, as well as machine learning techniques that can assist network management in adopting consistent measures to avoid interference, have been widely deployed. When it comes to the real network application, however, many of the most popular machine learning algorithms are deep learning, and they often focus on the engineering and selection elements; they are unable to successfully address the challenge of high-level login data separation. More segmentation jobs will be necessary as data sets grow in size, resulting in decreased accuracy.

1.2 Intrusion Detection Systems (IDS)

Intrusion is defined as any unauthorised activity that causes harm to a computer system. The goal of an IDS is to detect harmful network intrusion and computer action that a traditional firewall might miss. This is required to achieve high degrees of security against activities that threaten computer systems' reliability, security, or privacy. Signature based Intrusion Detection Systems (SIDS) and anomaly based Intrusion Detection Systems (AIDS) are the two types of IDS systems (AIDS).

1.2.1 Signature based IDS (SIDS)

A signature based IDS (SIDS), similar to antivirus software, examines all packets passing across a network and correlates them to such a collection of attack signatures or attributes of known malicious threats.

1.2.2 Anomaly based IDS (AIDS)

In order to monitors network and computer breaches and misuse, an anomaly based IDS observes significant moment and categorizes it as normal or abnormal. A considerable discrepancy between observed behavior and the model, referred to as an incursion, is described as an anomaly. This set of techniques is founded on the idea that hazardous behavior is distinct from ordinary user activity. Intrusions are instances of unusual consumer behavior that deviates from the standard.

2. Related Work

Alqahtani et al. [1] In this article, machine learning techniques were used to detect intrusions and evaluate the effectiveness of various cyber-security Experiments, including the Bayesian Classifier, Logistic Regression, Stochastic Decision Forest, Random Tree, Bayesian Network, Decision Table, and Artificial Neural Network. The performance measurements' precision, recall, f1 measure, and accuracy were examined using cyber-security datasets with a variety of cyber-attacks. The Random Decision Forest technique outperformed all of the authors' existing machine learning classification algorithms in terms of accuracy, recall, precision, and fscore. The author of this paper developed an information intrusion detection model that includes procedures including dataset exploration, data processing, and deep learning safety modelling. Subhash Waskle et al. [2] The following is a summary of how IDS works in this article on ML Approach to IDS: To identify signature packets, the IDS collects data or packets, preprocesses it, checks the signature, and then compares it to a signature database that is already available. The workings of anomaly detection are then explained: the data is preprocessed before patterns are generated. After that, it identifies outliers and generates reports. After that, data reduction techniques such as supervised, unsupervised, and semi-supervised procedures are discussed. There are also algorithms for feature categorization and feature extraction.

Feature selection: helps determine which set of features should be used to significantly improve result while reducing the number of errors. It cuts computer time and storage requirements greatly. PCA, IG, and GA are only a few examples. The wrapper technique, in which a classifier is utilized as a black box for ideal feature analysis, is described in the paper. Due to the computational cost of developing the classifier, such approaches may produce high speculation but may result in excessive complexity. Rows represent samples, while columns indicate features in the dataset. As a result, the removal aids in data dimension reduction while maintaining essential data.

After that, Clustering is discussed in further depth. Clustering groups data samples into packages of data, with data samples in each package being identical in some sense. Formula: TP+TN/TP+TN+FP+FN Accuracy: TP+TN/TP+TN+FP+FN TP/TP+FN attack detection rate FP/FP+TN is the false alarm rate. It then compares the accuracy of all of the different feature classifiers and extraction techniques that were employed on the same dataset kDD.

Subhash Waskle et al. [3] The authors used supervised machine learning approaches to find abnormalities in this paper. The Random Forest technique is used to choose k attributes in this model. The k characteristics are then separated again to generate node d, etc until you reach the starting node. This technique is performed endlessly for the generation of random forests. PCA was also utilized by the authors in their model, which is required for categorization. This method treats all of the data as a dataset with a large number of characteristics and a high dimension. Because all of the data points are on the same axis, the authors used this method to reduce the dimension of the data. After that, eigenvector and value characteristics are assessed, and a matrix is generated. This matrix is then used to calculate the major component. This model has an accuracy of roughly 90% when compared to SVM and Naive Bayes.

Ahmed I. Saleh, et al. [4] The authors of this paper proposed a different based intrusion detection and prevention method that hierarchical structures integrate an abuse detection model and an outlier detection model in a discretization framework that uses standard supervised algorithms for anomaly detection, including such 1 class SVM as well as j48 DT algorithms. This study used the NSL-KDD statistical model, which is a modified version of the well-known KDD Cup 99 benchmark datasets.

The authors used DT for the training models since it has a high detection accuracy. On normal training datasets, 1 class SVM is employed after decomposition using DT. DT decomposes the dataset and applies a single class SVM to each deconstructed segment. The anomaly detection model employs known attack information in an indirect approach to improve its capabilities while building normal behavior profiles throughout the integration. The hybrid intrusion detection technique, according to the authors' methodology for constructing their model, might improve IDS detection accuracy for unidentified intrusions and detection speed by considerably decreasing the maximum computation time of the train and test processes.

Waskle. S, et al. [5] In his research, a supervised machine learning method known as a Multilayer perceptron neural network was used for intrusion detection instead of typical machine learning techniques and neural network-based techniques like Logistic Regression, Stochastic Forest, or Svm Classifiers. The model was built using the CICIDS2017 dataset, which contains intrusion threats and traffic which is reflective of contemporary network usage. A fully connected feed-forward artificial neural classifier is used in this system's cross perceptron model. For the 14 types of attacks and traffic, the classifier has 15 outputs. CICIDS2017 has become one of them, and it contains intrusion threats as well as traffic that is realistic of current network activity. A confusion matrix was used to retrieve many significant components after training, including True Positive (TP), which is the lot of attacks receive similar to threats, True Negative (TN), which is the amount of majority class labelled as regular traffic, and False Positive (FP), which is the string of attacks categorized as threats and the number of victims prophesied as regular traffic. The authors were able to surpass typical machine learning techniques with this methodology, achieving an intrusion detection performance of over 99 per cent and a lower detection rate of less than 0.7 per cent.

Faker et al. [6] Faker used the CIC-2017 and UNSW-NB15 datasets to study intrusion detection. To avoid model overfitting, this study does not incorporate connection information. They delete blank values and unnecessary traffic data to minimize data volume. They also normalize the values after converting string values to numeric ones. They create two copies of the data set if there is missing or unlimited data. Got an average of all absent and infinite data to begin. Second, any data that is either absent or endless should be removed. They employ two types of datasets to test their model. DNN, Stochastic Forest, and Xgboost Tree classification are examples of training techniques. Zhang Xueqin et al. [7] X. Zhang exceled in using Deep Forest to prevent unauthorized access. They use the P-ZigZag encoding approach to preprocess the samples, and then perform an inverse discrete cosine transform to the preprocessed data sources. CIC-2018 displays more real-time traffic on the network, both with and without threats. CIC-2018 was created by compiling traffic on a network and computer information for over 80 various functions.

Yagnik Rathod et al. [8] Yagnik Rathod presented Database Intrusion Detection via Transaction Signature. They have presented a method for gaining access to the project's website administrative system. Their technique is based on security policies that have been accepted by the database management system. Signature Activity is used in the database management system to detect malicious transactions. Because these methods rely on the official Commercial signature, all legal transaction signals must be sent. They suggest using a Web Site Access to give an extra sense of protection to the site as well as to detect dangerous behaviors as part of any web application. This could aid in improving the DBMS security procedure. 16 By intervening to correct the unjust deed, they can improve the accuracy of the response.

3. Comparative Analysis

The Table 1 summarizes the previous work of wireless network intrusion detection by using various techniques. It also includes the goals of the existing works.

Table 1. Methods for detecting wireless intrusion

Paper	Method	Goal	Year
Kasongo et al.[9]	A deep learning-based wireless IDS.	The suggested FFDNN system enhances performance when compared to existing techniques.	2019

Kasango et al. [10]	A wireless intrusion detection system that employs a wrapper-based extracting features module as well as a Feed-Forward Deep Neural Network.	The UNSW-NB15 sample has 22 variables, with binary and multiclass classification techniques achieving accuracy rates of 87 percent and 77 percent, respectively. The AWID sample has 26 attributes and had accuracy results of 99 percent and 99 percent.	2020
Riyaz et al. [11]	A novel intrusion detection system (IDS) that recognizes and detects intruders to provide data communication security in wireless network.	In terms of accuracy rate (98.8%), training time (0.57 seconds), and testing time, the suggested model surpassed the competition (0.26 s).	2020
Satam et al.[12]	WIDS is a Wi-Fi (IEEE 802.11) protocol anomaly-based intrusion authentication protocol.	The suggested method detected wifi protocol assaults with a low error rate of positive (0.0174) and a variably low error rate of negatives for different threats.	2021
Singh et al. [13]	Deep learning techniques are used to create a novel algorithm.	In the NSL-KDD datasets, the accuracy rates for binary and multiclass classification were 97.32 percent and 97.47 percent, respectively.	2021
Devan et al. [14]	To classify network incursion, an XGBoost-DNN model is presented, which employs the XGBoost feature selection/ extraction technique and (DNN).	The NSL-KDD sample was used. Other approaches such as regression model, Bayesian network, and SVM were used to compare the results. Deep NN has a classification accuracy of 97 percent, which is higher than previous models, according to the data.	2020
Bedi et al. [15]	A 2-layer Enhanced Siam-IDS technique is employed to solve the problem of unbalanced class.	For both the CIDDS-001 and NSL-KDD samples, I-SiamIDS shows significant improvements in recalls, efficiency, f1 measure, precision, and Auc when compared to previous studies.	2021

4. Conclusion

This research reviews the literature for Wireless Network Intrusion Detection utilizing various machine learning, deep learning classification methods, and detection algorithms. From this study, it is observed that Convolutional Neural Network gives better results for the classification tasks. In future work, these techniques would be proposed for wireless network intrusion detection after applying the feature extraction method by using the univariate feature selection method.

References

[1] Alqahtani, Hamed, et al. "Cyber intrusion detection using machine learning classification techniques." International Conference on Computing Science, Communication and Security. Springer, Singapore, 2020.

[2] Mr. Subhash Waskle, Mr. Lokesh Parashar and Mr. Upendra Singh “Intrusion Detection System Using PCA with Random Forest Approach” Proceedings of the International Conference on Electronics and Sustainable Communication Systems (ICESC 2020).

[3] Subhash Waskle, Lokesh Parashar and Upendra Singh “Intrusion Detection System Using PCA with Random Forest Approach” International Conference on Electronics and Sustainable Communication Systems (ICESC),2020.

[4] Ahmed I. Saleh, Fatma M. Talaat and Labib M. Labib “A hybrid intrusion detection system (HIDS) based on prioritized knearest neighbors and optimized SVM classifiers” Artificial Intelligence Review volume 51, 2019.

[5] Waskle, S., Parashar, L., & Singh, U. (2020). Intrusion Detection System Using PCA with Random Forest Approach. 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC). doi:10.1109/icesc48915.2020.9155656

[6] Faker, Osama & Dogdu, Erdogan, “Intrusion Detection Using Big Data and Deep Learning Techniques,” in Proceedings of the 2019 ACM Southeast Conference, pp. 86-93. 2019.

[7] Zhang Xueqin, Chen Jiahao, Zhou Yue, Han, Liangxiu, Lin Jiajun, “A Multiple-layer Representation Learning Model for Network-Based Attack Detection,” IEEE Access. pp. 1-1. 10.1109/ACCESS.2019.2927465, 2019.

[8] Yagnik Rathod, Prof. M.B. Chaudhari, Prof. G.B. Jethava “Database Intrusion Detection by Transaction Signature”, IEEE, 2012.

[9] S. M. Kasongo and Y. Sun, ‘‘A deep learning method with wrapper based feature extraction for wireless intrusion detection system,’’ Comput. Secur., vol. 92, May 2020, Art. no. 101752.

[10] S. M. Kasongo and Y. Sun, ``A deep learning method with wrapper based feature extraction for wireless intrusion detection system," Comput.Secur., vol. 92, May 2020, Art. no. 101752.

[11] B. Riyaz and S. Ganapathy, ``A deep learning approach for effective intrusion detection in wireless networks using CNN," Soft Comput., vol. 24, no. 22, pp. 17265_17278, Nov. 2020.

[12] P. Satam and S. Hariri, ``WIDS: An anomaly based intrusion detectionsystem for Wi-Fi (IEEE 802.11) protocol," IEEE Trans. Netw. Service Manage., vol. 18, no. 1, pp. 1077_1091, Mar. 2021.

[13] N. B. Singh, M. M. Singh, A. Sarkar, and J. K. Mandal, ``A novel wide& deep transfer learning stacked GRU framework for network intrusion detection," J. Inf. Secur. Appl., vol. 61, Sep. 2021, Art. no. 102899.

[14] P. Devan and N. Khare, ``An ef_cient XGBoost_DNN-based classification model for network intrusion detection system," Neural Comput.Appl., vol. 32, pp. 12499_12514, Jan. 2020.

[15] P. Bedi, N. Gupta, and V. Jindal, ``I-SiamIDS: An improved siam-IDS for handling class imbalance in network-based intrusion detection systems,"Int. J. Speech Technol., vol. 51, no. 2, pp. 1133_1151, Feb. 2021.

Author's biography

S. Maheswari received her B.E degree in Computer Science and Engineering from the University College of Engineering, Trichy and pursuing M.E degree in Computer Science and Engineering from Government College of Technology, Coimbatore. Her research interests include Information Security, Machine Learning, Image Processing.

J. C. Miraclin Joyce Pamila is Professor & Head of the Department of Department of Computer Science and Engineering, Government College of Technology, Coimbatore. She teaches and guides students at both under graduate and Postgraduate levels. She is a life member of ISTE and currently takes up research in the area of Machine Learning, Data Analytics and Natural Language Processing.