INTERNATIONAL JOURNAL OF
RECENT TRENDS IN MULTIDISCIPLINARY RESEARCH
https://www.doi.org/10.59256/ijrtmr.20240404004
July-August 2024, Vol 4(04), 19-22.
ISSN No: 2583-0368

A Deep Learning Approach for Intrusion Detection Systems –A Review

Sivakumar Nagarajan
Technical Architect, I & I Software Inc, 2571 Baglyos Circle, Suite B-32, Bethlehem, Pennsylvania, USA.

OPEN ACCESS
Article Citation:
Sivakumar Nagarajan, "A Deep Learning Approach for Intrusion Detection Systems – A Review", International Journal of Recent Trends in Multidisciplinary Research, July-August 2024, Vol 4(04), 19-22.

Abstract: Throughout the past decade, the researchers have developed many a number of intrusion detection systems. Some of these were developed to work on host based and some were network-based intrusion detection systems. The presented systems are combination of HIDS/NIDS and signature/anomaly based, or hybrid systems. The intrusion detection system is an intelligence system known as computational intelligence. The main goal of computational intelligence is to provide solutions to complex real-world problems. An effective IDS must use more than standard mathematical techniques and conventional analysis methods combined with soft computing techniques to synergistically create a more robust IDS. In this paper, we review three important areas of research that have significant implication for proposed framework. First, we review existing shallow learning intrusion detection systems both in machine and deep learning. In the second section, we review existing hybrid intrusion detection systems. Finally, we provide some findings and analysis of the literature studies.

Key Words: Machine Learnings, NIDS, Detection systems, IDS

1. Introduction

In the present scenario, significant advancement of computer systems completely changed our daily lives and made our existence dependant on them. From small companies to large enterprises, individuals to government agencies, many of their activities are performed through network services. With the tremendous growth of the use of Internet, our computer systems are exposed to elevate high amount of threats. Any vulnerability in the network devices and computing platforms can expose the network system under various attacks and may lead to catastrophic consequences. However, it is a nightmare for organizations and corporation's security managers to prevent their networks from being attacked and to preserve their secretes and sensitive information of their customers from leaking out. Organizations should focus on innovative techniques to assist human analyst when dealing with surveillance, prevention, detection and response to cyber security incidents and potential attacks/intrusions on their network. The problem of how to effectively detect and prevent the intrusions in time is very challenging. Conventional network intrusion detection systems (NIDS) are signature or rule-based approaches that have not been adequate for the fast-growing network and unable to deal with attacks of their growing volume, complexity and deflation. The firewall is built to protect the entire network or systems from unauthorized access. The firewall and its variants have been shown that it could be easily bypassed by intruders, for instance, by using false source address. It also failed to detect so many attacks such as DoS and DDoS. To overcome the drawbacks of existing conventional security schemes, a new security mechanism has come to existence called "Intrusion Detection System" (IDS).

2. Literature Review

This section described the literature study of single, hybrid and ensemble intrusion detection techniques with machine learning methodology. Many of the researchers perform hybrid IDS to obtain good performance in terms of accuracy and detection rate.

Yassine Maleh et al. proposed a light weight IDS for Wireless Sensor Network (WSNs). This model uses anomaly detection based on support vector machine (SVM) algorithm and set of signature rules to detect abnormal behaviour. This methodology is integrated in a cluster-based topology, to reduce the communication cost. Simulation results show that the model could recognise anomalies effectively with high detection rate and with low false alarm rate.

Levent Koc et al. Proposed Hidden naïve Bayes data mining model that can be applied to intrusion detection problems which are affected by high dimensionality, high related features and high network data quantity. In this experiment they used two prominent variation techniques such as entropy minimization discretization and proportional k- interval discretization for good performance of implemented method on KDD99 dataset. The model includes feature selection model based on three filter methods such as correlation-based, consistency- based, and INTERACT feature selection methods. These techniques help to obtain good results in proposed method on KDD'99 datasets. Proposed hybrid model that integrated SVM and Ant Colony Optimization (ACO) which is used for distributed intrusion detection system (DIDS). ACO used for clustering and classification of attacks performed by SVM. This model provided higher detection rate and fast running time in real time environment. The hybrid model outperforms using only SVM or only ACO algorithm used for IDS.

3. Intrusion Detection Methodologies

IDS technologies use several methodologies to recognise intrusions. The major categories of detection methodologies are signature-based intrusion detection and anomaly-based intrusion detection. Most of IDS technologies use various detection methodologies, either separately or used together to detect intrusions accurately.

Signature-based Detection

Signature-based detection is also called as misuse-based detection; it retains the database of signatures of known attacks. A signature is a pattern that corresponds to a known threat. Upon collecting data from the audit unit, it matches the data against the database and alerts can be activated if a match is identified. These patterns describe suspect, collection of sequence of activities or operations that can be possibly harmful and stored in database. The primary benefit of this system is that it easily develops patterns and signatures and understands the network behaviour if familiar. It is more effective to maintain the attacks whose patterns are already preserved in the database.

Anomaly-based Detection

The other type of method is anomaly-based IDS, also usually known as behaviour-based IDS. These systems learn the regular behaviour of users, hosts, network connections or applications, i.e., hold the signature of the normal behaviour, instead of holding the signatures of known attacks. The profiles of behaviors are generated by observing the characteristics of typical activity over a period of time. Profiles can be generated for many behavioral attributes, as with the number of mails sent by a user, number of failed login attempts for host and level of processor usage in a given period of time. An Intrusion Detection System may also identify malicious behavior as it constructs its original profiles. Any deviation from the normal behaviour is deemed as suspicious and an alarm is activated. These systems operate under the assumption that any abnormal behaviour or activity differs significantly from the normal behaviour. By definition, these mechanisms are good at detecting zero-day attacks.

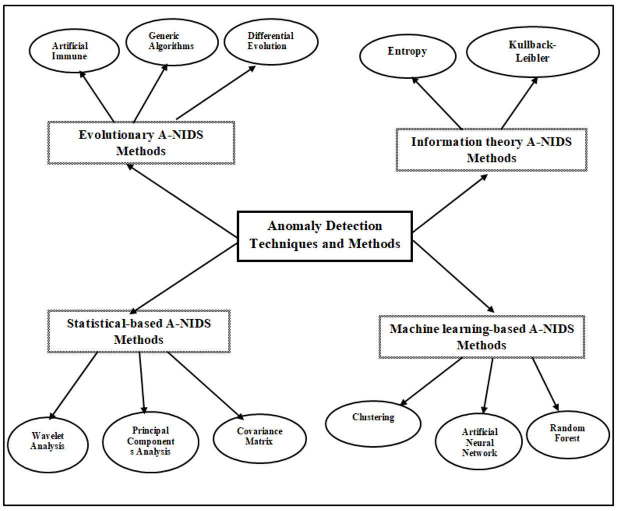
Artificial Immune
Generic Algorithms
Differential Evolution
Entropy
Kullback-Leibler
Evolutionary A-NIDS Methods
Information theory A-NIDS Methods
Anomaly Detection Techniques and Methods
Statistical-based A-NIDS Methods
Machine learning-based A-NIDS Methods
Wavelet Analysis
Principal Components Analysis
Covariance Matrix
Clustering
Artificial Neural Network
Random Forest

Figure 1 Anomaly detection techniques and methods

Hybrid based detection

To enhance the advantage and to eliminate the disadvantages of both methods, we need to combine the signature and anomaly-based methods. This method increases the detection rate of zero day attacks and manages to decrease the amount of false alarms. A survey observed that no system was purely signature or anomaly based, IDSs are generally deployed as hybrid of both signature and anomaly setup.

4. Problem Statement

The main target is anomaly-based intrusion detection systems and it makes major contribution in the region of anomaly-based IDS. It is well known that anomaly-based IDS endures from the high rate of false alarms. Ongoing efforts are being made to minimize the high false positive rate. We believe that intrusion detection is a data analysis processing and can be treated as a problem of classifying data properly. Based on this point, it can also be noticed that any classification scheme is good as the data provided to it as input has more clean data; precise results are likely to be obtained. From anomaly-based perspective, it signifies that if we can extract features that distinguish normal data from abnormal one appropriately, false positive rate can be minimized to a significant extent. Therefore, in this work, we explore the techniques which encourage in the process of distinguishing normal data from abnormal ones.

5. Artificial Intelligence

Artificial intelligence (AI) is an extensive area of computer science that makes machines function like a human brain. It is used to resolve complexity in problems to clarify using conventional computational techniques. Artificial Intelligence concept was formed in 1956, when computer scientists set out to determine if machines could "think" like humans. It was defined by Marvin Minksy as “the science of making machines do things that would require intelligence if done by human” [4]. Another equivalent definition for artificial intelligence, defined by Chollet [5] as “the effort to automate intellectual tasks normally performed by humans”. Artificial intelligence thus includes not only artificial learning and deep learning sub-fields, but also many strategies to enable the goal of automating intelligent activities usually done by humans. These strategies perform well for resolving well-defined logical activities such as playing games; however, they are not well equipped to deal with more complicated tasks, such as image classification and language translation. As a consequence, a modern approach to artificial intelligence, called machine learning, has gained popularity over existing methods to symbolic AI.

The field of intrusion detection is a region where current techniques mostly depend on human-programmed rules. Although there is space for these current intrusion prevention technologies, and they do well to implement unique parameters and block known malicious signatures. There are challenges in being able to adapt to modern, unseen threats that do not fall under the rigid set of rules that are described.

Supervised Learning

In supervised learning, predefined labelled dataset has been provided before training the algorithms. The primary objective of supervised learning is to learn how to assign best knowledge labels to new instances of the same activity that are not visible previously. As shown in the figure, a supervised machine learning algorithm consists of three components: a learning module, a model, and a classification module. A model based on a labelled training set is formulated by the learning module. The model comprises of a function that is formed by the learning module, and includes set of associative rules. These rules apply to an unlabeled and predict labels of the test samples. The estimation of the test sample labels is done with help of classification module.

Supervised learning models need to find the mapping function to map the input variable (X) with the output variable (Y). From the given set of instances X ∈{*x1, x2, x3,……, xn*} in which each instances xi may have one or more feature followed by the labels of each instance Y ∈{*y1, y2, y3, ……,yn*}. the leaning algorithm try to find a relationship inputs (instances) and outputs (labels) that represented by function h(x)=y. this function called as hypothesis and it will be used to predicting the output label of a given new instance *h(x1) = y1*.

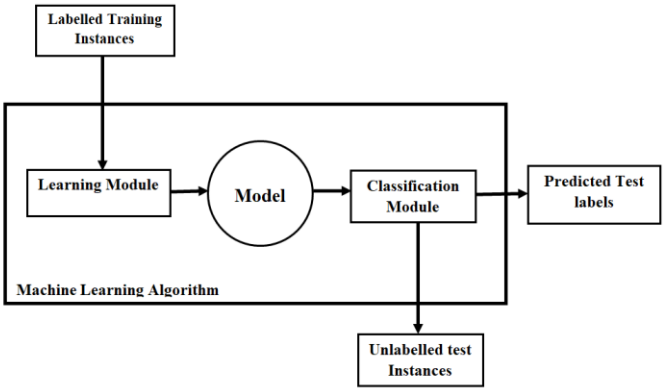
Labelled Training Instances
Learning Module
Model
Classification Module
Predicted Test labels
Machine Learning Algorithm
Unlabelled test Instances

Figure 2: Supervised machine learning, based on (Hendrickx, 2005)

6.Conclusion

It is concluded that most of the researchers used classifiers and cluster methods as an Intrusion Detection System. Later, feature extraction methods are used to extract the important features, along with classifiers. In this concerned area, many researchers used hybrid classifiers for Intrusion Detection, which are a combination of feature extracted, clustering, and classification methods.

After proper analysis of the above-discussed literature and techniques, we concluded that;

- hybrid classifier technique with feature selection methods. These methods attained higher accuracy compared with single classifiers. On the other hand, small amount of work is done on ensemble classifiers.
- Deep learning algorithms are suitable for feature learning and classification with huge amount of datasets. From the findings and analysis of literature study we conclude that, employing a feature reduction method is essential to reduce the computational cost and increase the classifier performance. Feature selection and feature extraction are having advantages and useful to detect attacks with classifiers, which make it hard with a single classifier method to implement. It’s recommended to use feature extraction followed by feature selection as a hybrid approach to increase the accuracy of intrusion detection.

Reference

1. *Wu, P. Deep learning for network intrusion detection: Attack recognition with computational intelligence. Master's Thesis, University of New South Wales, Sydney, NSW, Australia, 2020*
2. *Mighan, S.N.; Kahani, M. A novel scalable intrusion detection system based on deep learning. Int. J. Inf. Secur. 2021, 20, 387–403.*
3. *Liu, H.; Lang, B. Machine learning and deep learning methods for intrusion detection systems: A survey. Appl. Sci. 2019, 9, 4396.*
4. *Farhan, R.I.; Maolood, A.T.; Hassan, N.F. Optimized deep learning with binary PSO for intrusion detection on CSE-CIC-IDS2018 dataset. J. Al-Qadisiyah Comput. Sci. Math. 2020, 12, 16–27.*
5. *Bamasag, O.; Alsaeedi, A.; Munshi, A.; Alghazzawi, D.; Alshehri, S.; Jamjoom, A. Real-time DDoS flood attack monitoring and detection (RT-AMD) model for cloud computing. PeerJ Comput. Sci. 2022, 7, e814.*
6. *Bhardwaj, A.; Mangat, V.; Vig, R. Hyperband tuned deep neural network with well posed stacked sparse autoencoder for detection of DDoS attacks in cloud. IEEE Access 2020, 8, 181916–181929*
7. *Khraisat, A.; Gondal, I.; Vamplew, P.; Kamruzzaman, J.; Alazab, A. Hybrid intrusion detection system based on the stacking ensemble of c5 decision tree classifier and one class support vector machine. Electronics 2020, 9, 173.*