WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Intrusion Detection for Real-Time Network Datasets Using Machine Learning

Monika .D.Rokade, Dr.Prasad Upeddi

Research Scholar, JJTU, Jhunjhunu, Rajasthan
Guide,JJTU ,Jhunjhunu, Rajasthan

A B S T R A C T

In today's world, computer network and virtual machine security is critical. For network security or to restrict unwanted access by internal or external users, many designs have been suggested. Various current systems have been built to identify harmful activity on victim computers; sometimes, an external user develops malicious behavior and gains unauthorized access to victim machines using such a behavior system, which is referred to as malicious activity or intruder. To identify actions in real-time network log audit data, a variety of machine learning and soft computing algorithms have been developed. The most often used data sets to identify the Intruder on benchmark data sets are KKDDCUP99 and NLSKDD. In this study, we suggest employing machine learning methods to identify intrusions. A signature with detection and anomaly-based detection are two separate strategies that have been suggested. SVM, Nave Bayes, and ANN algorithms are shown with diverse data sets in the experimental study, as well as system performance in a real-time network context.

Keywords:Intrusion Detection System, Network security, Naïve Bayes, SVM, Artificial Neural Network, KDDCUP99.

1. INTRODUCTION

The IDS is only capable of detecting a single form of attack, such as a sample or unknown assault, a DoS attack, a U2R attack, or an R2L attack. It then installs a series of similar subsystems in order, one by one. This accomplishes two goals: In each sub-phase, only a restricted number of characteristics that identify a certain sort of assault may be taught. Second, the sub-size gadget is still small and useful.A typical disadvantage, similar to our technique, is that it increases communication overhead between modules. This may be readily avoided in our strategy by making each sub-phase totally independent of the other layers. As a result, such traits may occur in more than one sub-phase. Depending on the network's security policy, any sub-phase will simply stop an attack if it is discovered without a central decision maker. Various sub phases mainly function as filters inhibiting anomalous association as long as they are developed during a specific layer, giving speedy responsiveness to the incursion while also decreasing the analysis in subsequent stages.

2. LITERATURE SURVEY

According toRokade, Monika D., and Yogesh Kumar Sharma**.** Machine learning techniques are used to identify invaders. There are two types of identifiers: recognition and anomaly-based detection distinct methods that have been suggested. SVM, Nave Bayes, and ANN algorithms are shown utilizing various sets of data in the experimental study and performance of system in a network environment that is actual time.

Karuna S. Bhosale et al. deep neural network (DNN),As a kind of deep learning system, creating scalable and efficient IDSs to discover and distinguish inadvertent and unexpected cyber-attacks is being researched. The ongoing change in network functioning and the quick production of attacks need an evaluation of common occurrences developed throughout time using dynamic and static methodologies. This kind of research aids in the development of the most effective algorithm for predicting future cyber-attacks. On several public information test malware databases, a full review of the DNN trials and other sophisticated machine learning classifiers may be found. The KDDCup 99 dataset number of hidden layers used by the system is used to choose the optimum network parameters and modulation techniques for DNNs.

Chamou et al. ,Chamou et al. Because of that reason, the scientific community has gotten used to the complexity and improvement of intrusion detection systems' efficiency, a huge number of enterprises across the globe are being targeted and endangered by the continual appearance of new and evolving threats. This is a ground-breaking technology that uses deep learning models to evaluate suspicious activities in DDoS and malware cyber attacks. Because of the fast rise of online applications and their usage, most Internet users regard cybersecurity efficiency, data protection, and secure communication to be critical. Simultaneously, in the digital world of academia and business, particularly in small and medium-sized organisations (SMEs), growing vulnerability to more sophisticated security risks has been discovered across computers and internet networks, with economic ramifications.

According toWang, Jingyi, et al.an intrusion detection model with an attention mechanism based on a time-related deep learning technique. To begin, the system constructs a stacked sparse autoencoder (SSAE) to identify high-level representations of incursion data. To categorise traffic data, the system creates a two-layer bidirectional gated recurrent unit (BiGRU) network with an attention mechanism. The system is test the UNSW-NB15 benchmark dataset, the binary classification results show that employing high-dimensional sparse features generated by SSAE may greatly speed up classification progress. With a lower false alarm rate, higher accuracy, and less training and testing time, this model may successfully identify network intrusions and outperform previous comparable approaches.

According toAbdel-WahabFor exceptional situation network intrusion detection, machine learning model and deep learning model are evaluated. The study began with an overview of prior work in the subject of ML and DL IDS, followed by an overview of the datasets utilised in the studied literature. Furthermore, using the KDD-99 dataset, ML and DL models were evaluated, and performance results were given, contrasted, and debated. Finally, the authors suggest topics for further investigation that are crucial.

According toKrishna, Akhil, et al**.** As a result, this project aims to develop a Deep Learning-based intrusion detection/prevention system that can identify and prevent assaults like DOS, Probe, R2L, and U2R. When an intrusion occurs, it is identified with great accuracy using a Multi-Layer Perceptron Deep Learning model trained on the dataset kddcup99. Appropriate data from the network is collected and saved as a CSV file, then fed into the Deep Learning model developed to anticipate the assault in real-time, resulting in detection. The infiltration is stopped in the second step with the help of a background-running script. The script is designed to carry out the preventive phase by making informed decisions about which they should carry out prevention functions for various assaults. They may use data from the classification portion of the Multi-Layer Perceptron model to make a choice. A separate Intrusion Detection System and the Intrusion Prevention System are integrated as a single system in this article to accomplish the goal of quicker and more efficient intrusion detection and prevention activities.

According toRezaeipanah**,**A technique for integrating deep learning with observer learning in identifying intrusion patterns is described to improve computer network security. The characteristics of a deep neural network method that utilises combinations with representations of effective features are taught via observer learning. This technique is based on a supervision-based learning algorithm and a deep neural network that optimises the number of hidden layers and neurons in each layer depending on a threshold value. Experiments on the NSL-KDD data set indicate that the suggested approach outperforms MARS and DLNN algorithms with 97.64% accuracy.

3. PROPOSED SYSTEM

Using machine learning approaches, the suggested study methodology detected and prevented intrusions. The packet environments described block will conduct training, including packet selection for anomalous and remote monitoring. It will then submit a function collection for a certain packet action. If everything seems good, send it all together. Misbehaviour samples will be analysed for feature selection for different features in order to recognise specific assaults. The suggested system is divided into two phases; we used the network dataset for system training and testing—components of the framework. Figure 1 depicts the system's whole execution utilising stated algorithms. To build train modules and test them, several machine learning techniques were applied.

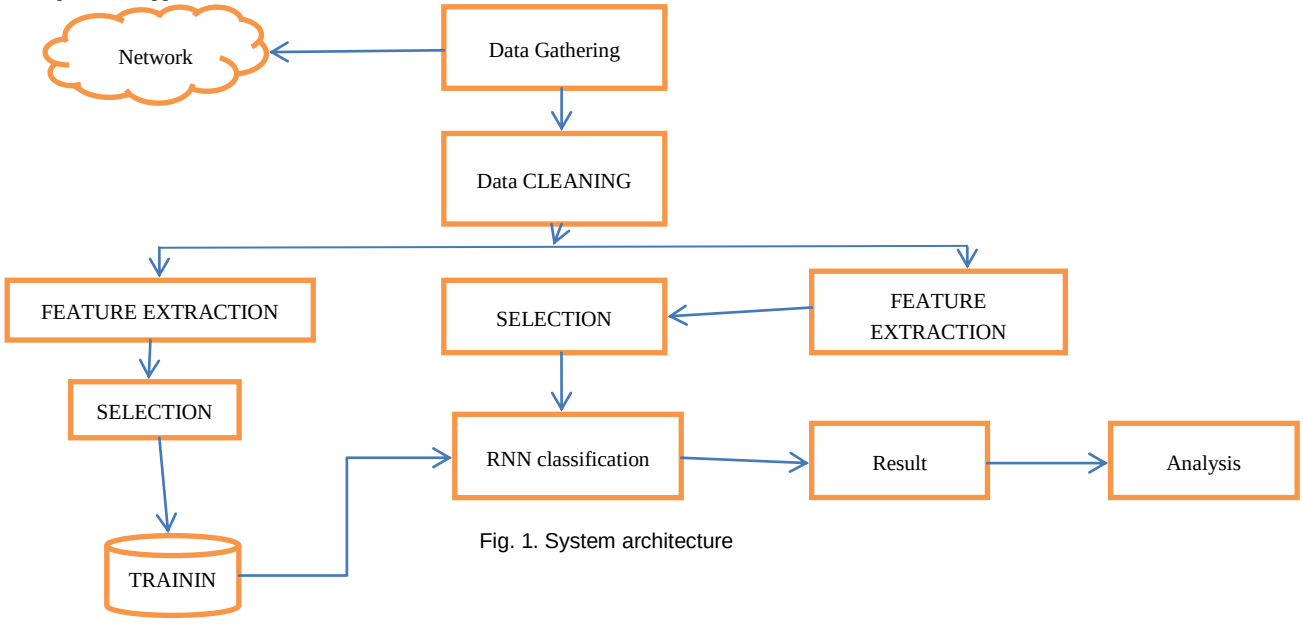

Fig. 1. System architecture

The experiments are based on the NSL KDD CUP 1999 dataset (KDD data set, 1999). The KDD CUP 1999 dataset is a ready-to-use and regulated version of the MIT Lincoln Laboratory's original 1998 DARPA intrusion detection assessment software. The stratification of the sampled consciousness is done. It has roughly 2 million association records as knowledge and around 5 million association records as coaching knowledge. The collection also includes a list of forty-one choices derived from each relationship, as well as a mark indicating whether records of associations are typical or specific forms of assault.

These choices are available for a wide variety of continuous, discrete, and symbolic variables, with radically diverse ranges falling into four categories: (1) The intrinsic choices of an association, including the important options for connecting individual transmission control protocols, make up the main class. The duration of the affiliation, the kind of protocol (TCP, UDP, etc.), and the network access (HTTP, telnet, etc.) area unit are all options. (2) The payload of the initial transmission control protocol packets is not determined by content selections inside an association reported by the Domain Data Area Unit, such as the aggregate of unsuccessful login attempts. (3) Constant host options assess specified connections with a continuous destination host throughout the previous 2 seconds due to the present relationship and measure protocol behaviour, operation, and other information.

(4) Connections with the same service as the current link during the previous two seconds are examined by similar service choices.

4. RESULTS AND DISCUSSION

We compute the system's confusion matrix once it has been successfully implemented. The classification using SVM techniques is shown in Tables 1 and 2. Figure 2 shows the classification performance of data collected by KDDCUP using the density-based approach of the machine learning algorithm programme Figure 3 shows how several approaches such as the RNN algorithm were used to classify and forecast the precision of the system.

TABLE I. CONFUSION MATRIX CALCULATION USING SVMFOR CLASSIFICATION

Class	Normal	Attack
Normal	1760	19
Attack	9	1640
	1769	1659

TABLE II. CONFUSION MATRIX CALCULATION USING NB FORCLASSIFICATION

Class	Normal	Attack
Normal	1830	1830
Attack	169	1202
	1999	1429

According to both experiment assessments, SVM outperforms the NB method in terms of classification accuracy (see Figure 3). According to the results of the preceding experiment, the system generates improved accuracy for trust calculation in the IoT in-service context. The whole study is guided by a set of simulated environmental settings and a mix of machine learning methods. With regards to machine learning techniques, many calculation parameters have been employed cluster differentiation.

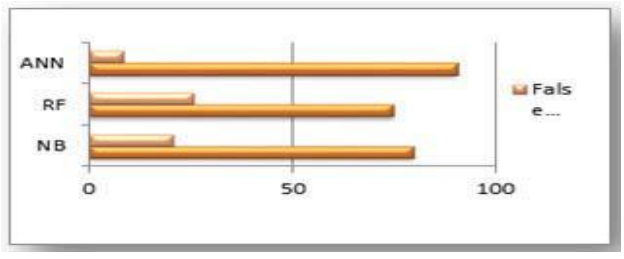

Fig. 2. Detection accuracy for KD D : CUP99 dataset using machine learning

The above figure 2 Shows accuracy of kddCup 99 results classification, with five different classes. Average software output is around the algorithm for the machine learning 88.00% for all classes.

5. CONCLUSION

This work presented a deep learning-based SVM-IDS technique for recommending an efficient ID system. The synthetic-based intrusion dataset NSL-KDD was used to assess anomaly detection accuracy. We want to use deep learning to build IDS in the cloud environment in the future. We also examine and compare several deep learning approaches, such as. The software mainly functions as artificial intelligence and conditioning algorithm to find the unknown occurrences during the data check by using NB ANN, RF, and SVM on the NSL-KDD dataset to detect network intrusions. Improved classification and high-class detection are possible because to the efficient rule structure. Several trials employed experimental analysis to evaluate the

algorithm's effectiveness using a variety of tests, and we came to the conclusion that we were getting adequate results.

References

[1] Rokade Monika D., and Yogesh Kumar Sharma. "MLIDS: A Machine Learning Approach for Intrusion Detection for Real Time Network Dataset." 2021 International Conference on Emerging Smart Computing and Informatics (ESCI).IEEE, 2021.

[2] Chamou, Dimitra, et al. "Intrusion Detection System Based onNetwork Traffic Using Deep Neural Networks." 2019 IEEE 24thInternational Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD). IEEE, 2019.

[3] Mohammed Anbar, RosniAbdulah, Izan H. Hasbullah, Yung- Wey Chong; Omar E. Elejla, "Comparative Performance Analysis of classification algorithm for Internal Intrusion Detection ", 2016 14thAnnual Conference on Privacy Security and Trust (PCT), Dec 12- 14,2016, Penang, Malaysia.

[4] Weiwei Chen, Fangang Kong, Feng Mei, GuiginYuan, Bo Li, "a novel unsupervised Anamoly detection Approach for Intrusion Detection System", 2017 IEEE 3rd International Conference on big data security on cloud, May 16-18,2017, Zhejiang, China.

[5] Anna L. Buczak, Erha n Guven, "A Survey of Data Mining and Machine Learning methods for cybersecurity intrusion detection", IEEE communication surveys and tutorials, vol. 18, Issue 2,2016.

[6] Manoj s. Koli, Manik K. Chavan, "An Advanced method for detection of botnet traffic using Interhnal Intrusion Detection", 2017 International Conference on (ICICCT), March 10-11, 2017, Sangli,India.

[7] ThabetKacem, DumindaWijesekera, Paulo Costa, Alexander Barreto, "An ADS-B Intrusion Detection System", 2016 IEEE on ISPA, 2016, Fairfax, Virginia.

[8] Shengyi Pan, Thomas Morris, UttamAdhikari, "Developing a Hybrid Intrusion Detection System using Data Mining for power system",
IEEE Transactions on, vol. 6, issues. 6, Nov. 2015.

[9] Mehdi Ezzarii, Hamid Elghazi, Hassan El Ghazi, Tayeb Sadiki, "Epigenetic Algorithm for performing Intrusion Detection System", 2016 International Conference on ACOSIS, Oct17- 19,2016, Rabat, Morocco.

[10] BOROLE, Prajakta; SHARMA, Yogesh Kumar; NEMADE, Santosh.6G Network Access and Edge-Assisted Congestion Rule Mechanism using Software-Defined Networking. International Journal of Future Generation Communication and Networking, 2020, 13.1s: 107-112.