ISSN: 0976-0172

Journal of Bioscience And Technology
www.jbstonline.com

A Neural Network-Based Approach for Identifying Cyber security Risks Through Event Profile Analysis

Mr.K.Naresh[1]., G.Prathyusha[2]

1 Assistant Professor, Department of CSE, Malla Reddy College of Engineering for Women., Maisammaguda., Medchal., TS, India

2, B.Tech CSE (19RG1A0516),

Malla Reddy College of Engineering for Women., Maisammaguda., Medchal., TS, India

Article Info

Received: 28-05-2022 *Revised: 16-07-2022* *Accepted: 27-07-2022*

Abstract—

Developing an automated method for detecting cyber threats is one of the main difficulties facing cyber security. In this study, we describe an artificial neural network-based method for detecting cyber threats. The suggested approach improves cyber-threat identification by converting a large number of gathered security events into unique event profiles and using a deep learning-based detection algorithm. In order to accomplish this task, we created an AI-SIEM system that combines several artificial neural network techniques, such as CNN, LSTM, and FCNN, with event profiling for data preparation. The approach helps security analysts react quickly to cyber threats by focusing on differentiating between real positive and false positive signals. The authors of this work used two benchmark datasets (NSLKDD and CICIDS2017) as well as two real-world datasets to conduct all of the experiments. We ran trials utilizing the five traditional machine-learning techniques (SVM, k-NN, RF, NB, and DT) to assess the performance comparison with current methodologies. As a consequence, the study's experimental findings confirm that our suggested approaches may be used as learning-based models for network intrusion detection and demonstrate that, despite being used in the real world, their performance surpasses that of traditional machine learning techniques.
SVM, k-NN, RF, NB, DT, AI-SIEM, FCNN, CNN, and LSTM are the index terms.

I Introduction

The development of artificial intelligence (AI) techniques has led to the improvement of learning-based methods for identifying cyber attacks, which have produced noteworthy findings in several research. However, it is still very difficult to defend IT systems against threats and bad activities in networks since cyber attacks are always changing. In order to establish trustworthy solutions, robust defenses and security considerations were given top importance due to many network intrusions and harmful actions. Traditionally, there have been two main methods for identifying network breaches and cyber threats. The company network has an intrusion prevention system (IPS) installed, which largely uses signature-based techniques to inspect network protocols and flows. It creates relevant intrusion alarms, also known as security events, and communicates the alerts it creates to another system, like SIEM. The collection and handling of IPS alerts has been the primary emphasis of security information and event management, or SIEM. Among the several security operations solutions available for analyzing the gathered security data, the SIEM is the most widely used and trustworthy option. Additionally, security analysts try to look into suspicious alerts by threshold and rules, and they use attack-related information to analyze connections between events in order to find malicious conduct. Due to the large volume of security data and the high false alarm rate of intelligent network assaults, it is still challenging to identify and detect breaches. For this reason, machine learning and artificial

ISSN: 0976-0172

Journal of Bioscience And Technology
www.jbstonline.com

intelligence algorithms for attack detection have received more attention in the most recent research in the area of intrusion detection. Developments in AI disciplines may help security analysts investigate network attacks more quickly and automatically. These learning-based techniques need using previous threat data to understand the attack model, then using the learned models to find incursions for unidentified cyber threats.

For analysts who need to quickly examine a lot of events, a learning-based approach designed to ascertain if an attack happened in a lot of data might be helpful. Information security solutions may be broadly classified into two types: machine learning-driven solutions and analyst-driven solutions. Analyst-driven solutions are based on rules that are established by analysts, who are security professionals. Meanwhile, new cyber threat detection may be enhanced by machine learning-driven technologies that identify uncommon or unusual behaviors. However, we found that the current learning-based techniques have four major drawbacks, despite the fact that they are helpful in identifying cyber attacks in systems and networks.

2 Review of the Literature on Deep Neural Network-Based Improved Network Anomaly Detection

Summary: The past ten years have seen an enormous rise in Internet applications, which has made information network security more important. An intrusion detection system is supposed to adapt to a constantly changing threat environment as the first line of defense for network infrastructure. Researchers in the fields of data mining and machine learning have developed a variety of supervised and unsupervised methods to reliably identify abnormalities. In the field of machine learning, deep learning uses a structure like to a neuron to accomplish learning tasks. Deep learning has revolutionized the way that learning tasks are approached by bringing about enormous advancements in a variety of fields, including computer vision, audio processing, and natural language processing, to mention a few. The sole use for this new technology that warrants investigation is in information security. This research looks at whether deep learning techniques are suitable for anomaly-based intrusion detection systems. In this study, we created models for anomaly detection based on several deep neural network architectures, such as recurrent neural networks, auto encoders, and convolutional neural networks.

These deep models were assessed using the two test data sets supplied by after being trained on the training data set.

The authors conducted every experiment in this study using a GPU-based test bench. Well-known classification approaches, such as extreme learning machine, nearest neighbor, decision-tree, random forest, support vector machine, naive-bays, and quadratic discriminate analysis, were used to create conventional machine learning-based intrusion detection models. Well-known classification criteria, such as receiver operating characteristics, area under the curve, precision-recall curve, mean average precision, and classification accuracy, were used to assess both deep learning and traditional machine learning models. Deep IDS model experimental findings demonstrated encouraging outcomes for practical use in anomaly detection systems.

3 Implementation Study

Traditionally, there have been two main methods for identifying network breaches and cyber threats. The company network has an intrusion prevention system (IPS) installed, which largely uses signature-based techniques to inspect network protocols and flows. It creates relevant intrusion alarms, also known as security events, and notifies another system—like SIEM—of the alerts it creates. The collection and handling of IPS alerts has been the primary emphasis of security information and event management, or SIEM. Among the many security operations solutions, the SIEM is the most popular and reliable option for analyzing the gathered security events. In addition, security analysts try to look into suspicious alerts based on policies and thresholds and find malicious activity by looking for patterns in the events and applying attack-related knowledge to analyze correlations between them.

Proposed Methodology

By grouping events together using a concurrency feature and establishing correlations between event sets in the data obtained, the suggested AI-SIEM system specifically comprises an event pattern extraction approach. Our event profiles may be used as succinct source data for different types of deep neural networks. Additionally, it makes it possible for the analyst to compare all of the data with long-term historical data in a timely and effective manner.

The developed artificial intelligence (AI)-based SIEM system's workflow and architecture. The three primary stages of the AI-SIEM system are the real-time threat detection phase, the learning engine based on artificial neural networks, and the data preparation phase. Through the transformation of raw data, the

ISSN: 0976-0172

Journal of Bioscience And Technology
www.jbstonline.com

system's first preprocessing step, known as event profiling, seeks to provide condensed inputs for different deep neural networks. The AI-SIEM system performs data preprocessing, data aggregation with parsing, data normalization using the TF-IDF technique, and event profiling in that order. As shown in Figure, each step produces event data sets, event vectors, and event profiles, respectively, and uses the result in the subsequent stage. This step comes before both the data learning stage and the process of converting raw security events into input data for the deep learning engine when the system is used to identify network breaches in real time. For modeling, the second AI-based learning engine uses three artificial neural networks. The preprocessed data are input into each of the three artificial neural networks (ANNs) for the data learning step, where each ANN learns to identify the best correct model. Lastly, each ANN model uses the trained model to automatically classify each security raw event in real-time threat detection. The dashboard provides security analysts with only verified genuine warnings, hence minimizing false ones.

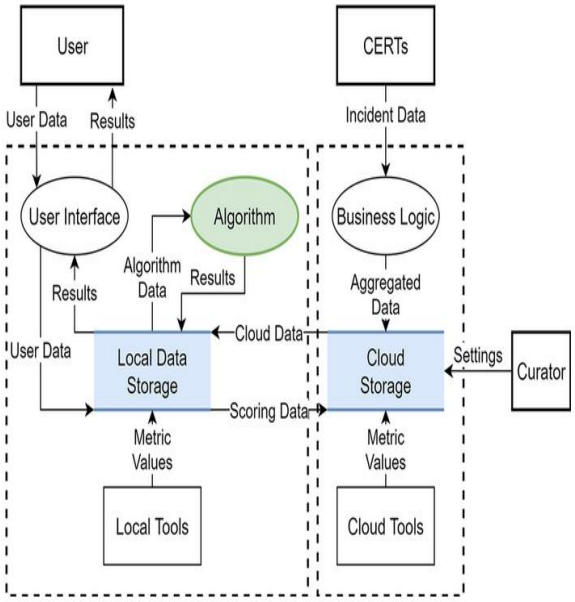

Fig 1: - flow of proposed system

Methodology

Uploading the train dataset; Running the TF-IDF preprocessing algorithm; Generating an event vector; Neural Network Profiling; Executing SVM, KNN, and Naive Bayes algorithms; Executing Decision Tree algorithms; Creating an accuracy comparison graph; Creating a precision comparison graph; Creating a recall comparison graph; and Creating an FMeasure comparison graph.

1) Data Parsing: To generate a raw data event model, this module parses an input dataset.

2) TF-IDF: We will use this module to transform unprocessed data into an event vector that includes both attack and normal signatures.

3) Event Profiling Stage: Based on event profiling, processed data will be divided into train and test models.

4) Deep Learning Neural Network Model: This module creates a training model by using CNN and LSTM algorithms to train and test data. The generated trained model will be used to compute FMeasure, Recall, Precision, and prediction score on test data. An algorithm that learns flawlessly will provide results with higher accuracy, and that model will be chosen to be used for attack detection on a real system.

The testing datasets we are using are quite large, and kdd_train will fail with an out of memory error during model construction. The CSV dataset is operating flawlessly, however it will take five to ten minutes to execute every algorithm. The remaining datasets may also be tested by scaling them down or executing them on a machine with a lot of settings.

4 Results and Evolution Metrics

Fig 2:_ Main screen

ISSN: 0976-0172

Journal of Bioscience And Technology
www.jbstonline.com

Fig 3: In above screen uploading 'kdd_train.csv' dataset and after upload will get below screen

Fig 4: _ a In above screen we can see dataset contains 9999 records and now click on 'Run Preprocessing

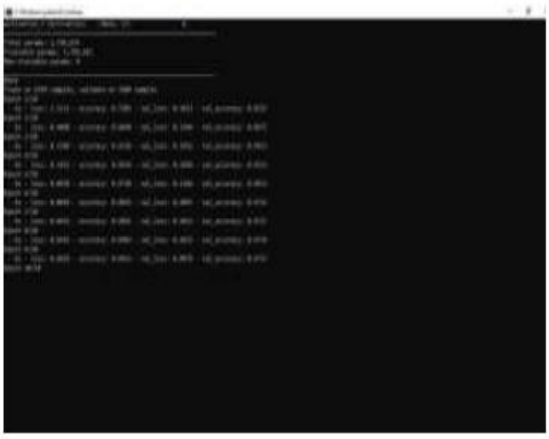

Fig 5: _ In the screen above, CNN likewise begins the first iteration with an accuracy of 0.72. After 10 iterations, we have filtered out an improved accuracy of 0.99, which we can multiply by 100 to get an accuracy of 99%. Thus, CNN is providing more accuracy than LSTM, and you can see the whole GUI screen below.

Fig 6:- From the following graph, which shows the name of the algorithm and its accuracy on the y-axis, we may infer that CNN and LSTM perform well. To see the graph below, click Precision Comparison Graph now.

Conclusion

We have presented the AISIEM system in this study, which makes use of artificial neural networks and event profiles. Condensing very huge amounts of data into event profiles and using deep learning-based detection techniques to improve cyber-threat detection capabilities are the innovative aspects of our study. By comparing long-term security data, the AI-SIEM system helps security analysts to respond to important security alarms quickly and effectively. It may also assist security analysts in quickly responding to cyber threats scattered over a multitude of security events by decreasing false positive alarms. We conducted a performance comparison utilizing two benchmark datasets (NSLKDD, CICIDS2017) and two real-world datasets to assess performance. Using well-known benchmark datasets and comparative experiments, we first demonstrated the applicability of our processes as one of the learning-based models for network intrusion detection. Second, we demonstrated encouraging findings from the assessment using two actual datasets, showing that our approach performed better in terms of accurate classifications than traditional machine learning techniques.

5 References

[1] S. Naseer, Y. Saleem, S. Khalid, M. K. Bashir, J. Han, M. M. Iqbal,K. Han, "Enhanced Network Anomaly Detection Based on DeepNeural

ISSN: 0976-0172

Journal of Bioscience And Technology
www.jbstonline.com

Networks,"IEEE Access, vol. 6, pp. 48231-48246, 2018.

[2] B. Zhang, G. Hu, Z. Zhou, Y. Zhang, P. Qiao, L. Chang, "NetworkIntrusion Detection Based on Directed Acyclic Graph and Belief RuleBase", ETRI Journal, vol. 39, no. 4, pp. 592-604, Aug. 2017

[3] W. Wang, Y. Sheng and J. Wang, "HAST-IDS: Learning hierarchicalspatialtemporal features using deep neural networks to improveintrusion detection,"IEEE Access, vol. 6, no. 99, pp. 1792-1806,2018.

[4] M. K. Hussein, N. Bin Zainal and A. N. Jaber, "Data security analysisfor DDoS defense of cloud based networks,"2015 IEEE StudentConference on Research and Development (SCOReD), KualaLumpur, 2015, pp. 305-310.

[5] S. Sandeep Sekharan, K. Kandasamy, "Profiling SIEM tools andcorrelation engines for security analytics,"In Proc. Int. Conf.Wireless Com., Signal Proce. and Net.(WiSPNET), 2017, pp. 717-721.

[6] N.Hubballiand V.Suryanarayanan,''False alarm minimization techniques in signature-based intrusion detection systems: Asurvey,'' Comput. Commun., vol. 49, pp. 1-17, Aug. 2014.

[7] A. Naser, M. A. Majid, M. F. Zolkipli and S. Anwar, "Trusting cloudcomputing for personal files,"2014 International Conference onInformation and Communication Technology Convergence (ICTC),Busan, 2014, pp. 488-489.