ISSN 2277-2685
IJESR/AUG 2022/ Vol-12/Issue-3/1-12
Mr.A.Madhu / International Journal of Engineering & Science Research

PACKET INSPECTION TO IDENTIFY NETWORK LAYER ATTACKS USING MACHINE LEARNING

Mr.A.Madhu,
Assistant Professor, Dept. of CSE,
Malla Reddy Engineering College (Autonomous), Secunderabad, Telangana State

ABSTRACT:

Intrusion detection can identify unknown attacks from network traffics and has been an effective means of network security. Nowadays, existing methods for network anomaly detection are usually based on traditional machine learning models, such as KNN, SVM, etc. Although these methods can obtain some outstanding features, they get a relatively low accuracy and rely heavily on manual design of traffic features, which has been obsolete in the age of big data. To solve the problems of low accuracy and feature engineering in intrusion detection, a traffic anomaly detection model BAT is proposed. The BAT model combines BLSTM (Bidirectional Long Short-term memory) and attention mechanism. Attention mechanism is used to screen the network flow vector composed of packet vectors generated by the BLSTM model, which can obtain the key features for network traffic classification. In addition, we adopt multiple convolutional layers to capture the local features of traffic data. As multiple convolutional layers are used to process data samples, we refer BAT model as BAT-MC. The softmax classifier is used for network traffic classification. The proposed end-to-end model does not use any feature engineering skills and can automatically learn the key features of the hierarchy. It can well describe the network traffic behavior and improve the ability of anomaly detection effectively. We test our model on a public benchmark dataset, and the experimental results demonstrate our model has better performance than other comparison methods.

Key words: BAT, MAC, CNN, DL, ML.

I INTRODUCTION

Intrusion detection plays an important part in ensuring network information security. Machine learning methods have been widely used in intrusion detection to identify malicious traffic.

*Corresponding Author

However, these methods belong to shallow learning and often emphasize feature engineering and selection. They have difficulty in features selection and cannot effectively solve the massive intrusion data classification problem, which leads to low recognition accuracy and high false alarm rate. In recent years, intrusion detection methods based on deep learning have been proposed successively.

1.2 PROBLEM DEFINITION

The existing methods for network anomaly detection are usually based on traditional machine learning models, such as KNN, SVM, etc. Although these methods can obtain some outstanding features, they get a relatively low accuracy and rely heavily on manual design of traffic features, which has been obsolete in the age of big data.

1.3 OBJECTIVE OF PROJECT

we adopt multiple convolutional layers to capture the local features of traffic data. As multiple convolutional layers are used to process data samples, we refer BAT model as BAT-MC. The softmax classifier is used for network traffic classification.

2.LITERATURE SURVEY

2.1 A Survey: Intrusion Detection Techniques for Internet of Things

AUTHORS:Sarika Choudhary and Nishtha Kesswani

The latest buzzword in internet technology nowadaysisthe Internet of Things. The Internet of Things (IoT) is an ever-growing network which will transform real-world objects into smart or intelligent virtual objects. IoT is a heterogeneous network in which devices with different protocols can connect with each other in order to exchange information. These days, human life depends upon the smart things and their activities. Therefore, implementing protected communications in the IoT network is a challenge. Since the IoT network is secured with authentication and encryption, but not secured against cyber-attacks, an Intrusion Detection System is needed. This research article focuses on IoT introduction, architecture, technologies,

attacks and IDS. The main objective of this article is to provide a general idea of the Internet of Things, various intrusion detection techniques, and security attacks associated with IoT.

2.2 Network intrusion detection

AUTHORS:B. Mukherjee, L.T. Heberlein and K.N. Levitt

Intrusion detection is a new, retrofit approach for providing a sense of security in existing computers and data networks, while allowing them to operate in their current "open" mode. The goal of intrusion detection is to identify unauthorized use, misuse, and abuse of computer systems by both system insiders and external penetrators. The intrusion detection problem is becoming a challenging task due to the proliferation of heterogeneous computer networks since the increased connectivity of computer systems gives greater access to outsiders and makes it easier for intruders to avoid identification. Intrusion detection systems (IDSs) are based on the beliefs that an intruder's behavior will be noticeably different from that of a legitimate user and that many unauthorized actions are detectable. Typically, IDSs employ statistical anomaly and rulebased misuse models in order to detect intrusions. A number of prototype IDSs have been developed at several institutions, and some of them have also been deployed on an experimental basis in operational systems. In the present paper, several host-based and network-based IDSs are surveyed, and the characteristics of the corresponding systems are identified. The host-based systems employ the host operating system's audit trails as the main source of input to detect intrusive activity, while most of the network-based IDSs build their detection mechanism on monitored network traffic, and some employ host audit trails as well. An outline of a statistical anomaly detection algorithm employed in a typical IDS is also included.

2.3 survey on sdn based network intrusion detection system using machine learning approaches

AUTHORS: N. Sultana, N. Chilamkurti, W. Peng, and R. Alhadad

Software Defined Networking Technology (SDN) provides a prospect to effectively detect and monitor network securityproblems ascribing to the emergence of the programmable features. Recently, Machine Learning (ML) approaches have beenimplemented in the SDN-based Network

Intrusion Detection Systems (NIDS) to protect computer networks and to overcomenetwork security issues. A stream of advanced machine learning approaches –the deep learning technology (DL) commences toemerge in the SDN context. In this survey, we reviewed various recent works on machine learning (ML) methods that leverageSDN to implement NIDS. More specifically, we evaluated the techniques of deep learning in developing SDN-based NIDS. Inthe meantime, in this survey, we covered tools that can be used to develop NIDS models in SDN environment. This survey isconcluded with a discussion of ongoing challenges in implementing NIDS using ML/DL and future works.

Existing system:

Most algorithms have been considered for use in the past. In [16], the authors make a summary of pattern matching algorithm in Intrusion Detection System: KMP algorithm, BM algorithm, BMH algorithm, BMHS algorithm, AC algorithm and AC-BM algorithm. Experiments show that the improved algorithm can accelerate the matching speed and has a good time performance. In [17], Naive approach, Knuth-MorrisPratt algorithm and RabinKarp Algorithm are compared in order to check which of them is most efficient in pattern/intrusion detection. Pcap files have been used as datasets in order to determine the efficiency of the algorithm by taking into consideration their running times respectively.

DISADVANTAGES OF EXISTING SYSTEM:

1. we are also facing various security threats. Network viruses, eavesdropping and malicious attacks are on the rise, causing network security to become the focus of attention of the society and government departments.

2. to identify various malicious network traffics, especially unexpected malicious network traffics, is a key problem that cannot be avoided.

PROPOSED SYSTEM:

The accuracy of the BAT-MC network can reach 84.25%, which is about 4.12% and 2.96% higher than the existing CNN and RNN model, respectively. The following are some of the key contributions and findings of our work:

1) We propose an end-to-end deep learning model BAT-MC that is composed of BLSTM and attention mechanism. BAT-MC can well solve the problem of intrusion detection and provide a new research method for intrusion detection.

2) We introduce the attention mechanism into the BLSTM model to highlight the key input. Attention mechanism conducts feature learning on sequential data composed of data package vectors. The obtained feature information is reasonable and accurate.

3) We compare the performance of BAT-MC with traditional deep learning methods, the BAT-MC model can extract information from each packet. By making full use of the structure information of network traffic, the BAT-MC model can capture features more comprehensively.

4) We evaluate our proposed network with a real NSL-KDD dataset. The experimental results show that the performance of BAT-MC is better than the traditional methods.

ADVANTAGES OF PROPOSED SYSTEM:

1.The BAT-MC model consists of five components, including the input layer, multiple convolutional Layers, BSLTM layer, attention layer and output layer, from bottom to top.

2. At the input layer, BAT-MC model converts each traffic byte into a one-hot data format. Each traffic byte is encoded as an n-dimensional vector. After traffic byte is converted into a numerical form, we perform normalization operations.

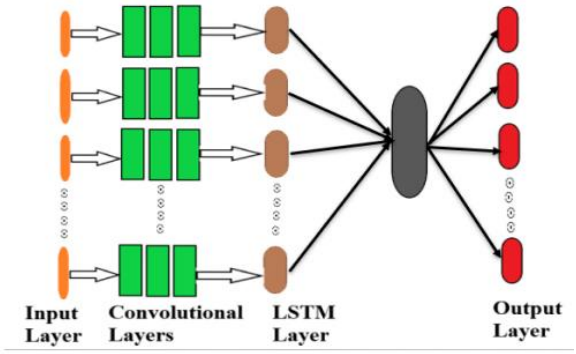
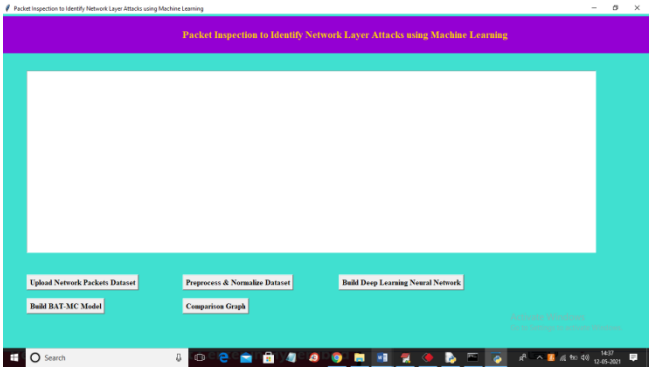


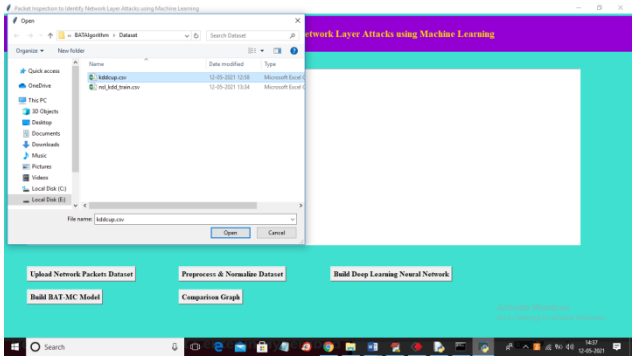
Fig.1. SYSTEM ARCHITECTURE

METHODOLOGY

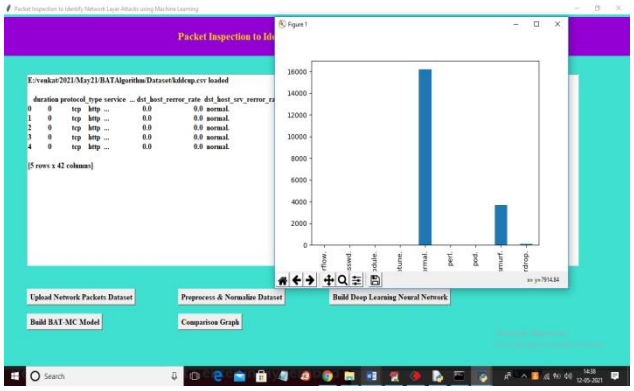
To run project double click on 'run.bat' file to get below screen



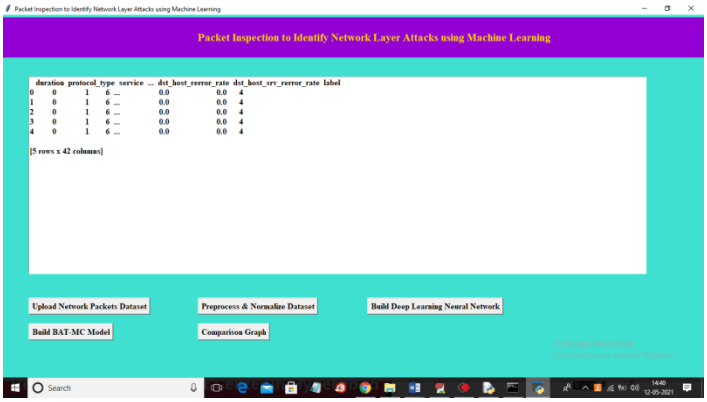
In above screen click on 'Upload Network Packets Dataset' button to upload dataset and to get below screen



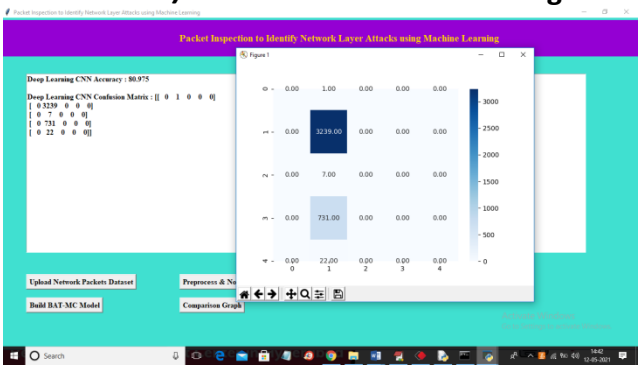
In above screen selecting and uploading 'kddcup.csv' file and then click on 'Open' button to load dataset and to get below screen



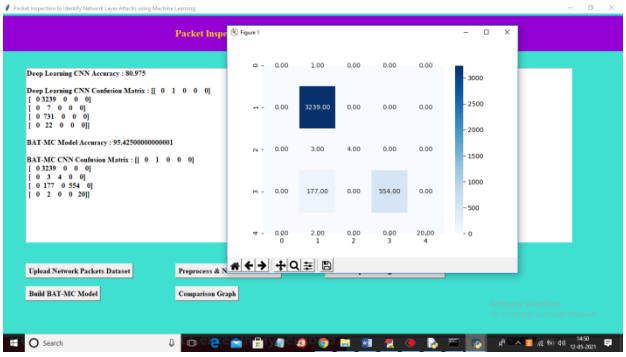
In above screen in text area we can see dataset loaded and we can see data contains alpha numeric data and ML algorithms accept only numeric values so we need to preprocess and normalize them and in graph we can see different attack names in x-axis and total attack types on y-axis and now close above graph and then click on 'Preprocess & Normalize Dataset' button to normalize data



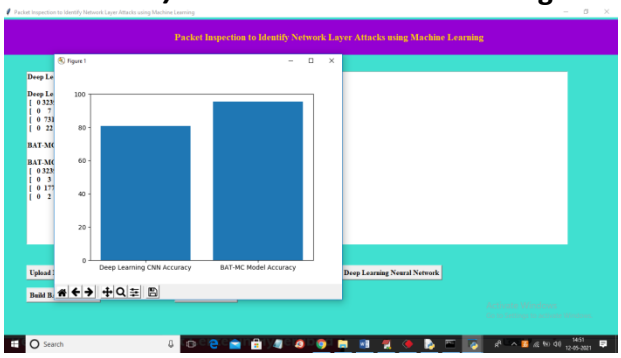
In above screen we can see dataset converted to numeric values by assigning ID's to each unique non-numeric data and now dataset is ready and now click on 'Build Dee Learning Neural Network' button to train CNN above dataset and then calculate prediction accuracy



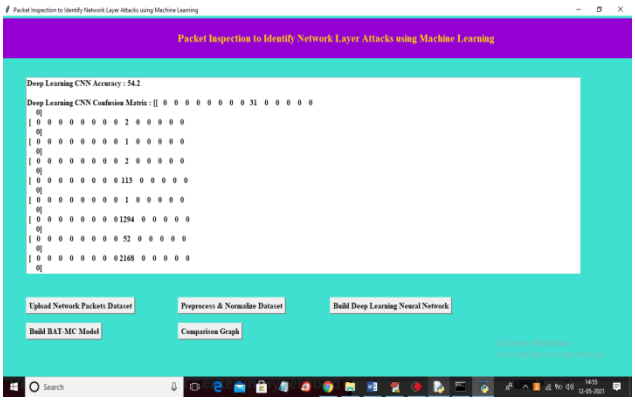
In above screen we can see CNN algorithm got 80% accuracy and in confusion matrix we can see total 5 different attacks are found and in confusion matrix we can see which attack predicted how many times. For example attack 2 predicted 3239 times in entire test data. Now close above graph and then click on 'Build BAT-MC Model' to train above dataset with BLSTM algorithm and then calculate prediction accuracy on test data.



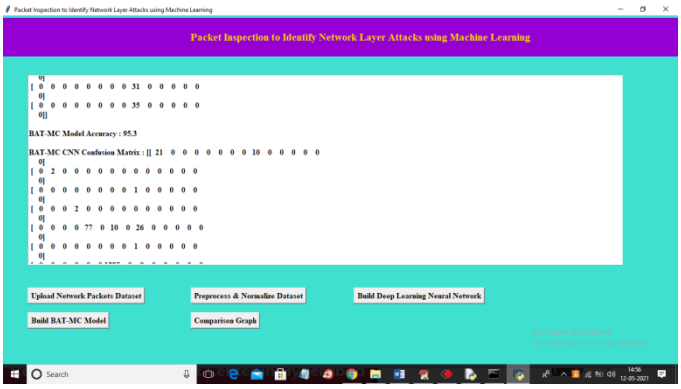
In above screen BAT-MC model generated and its prediction accuracy is 95 and now close above graph and then click on 'Comparison Graph' button to get below graph



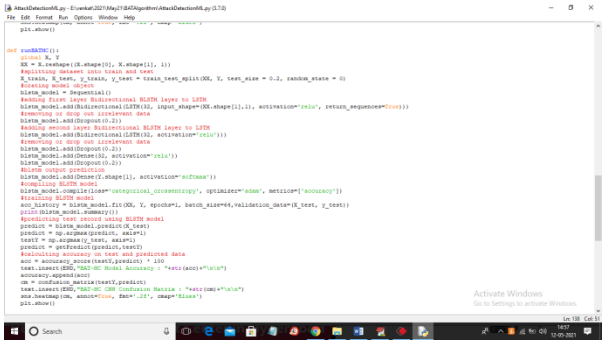
In above graph x-axis represents algorithm name and y-axis represents accuracy and in both algorithm BAT-MC model is giving better accuracy. Similarly you can upload other dataset and can build CNN and BAT-MC model. In below screen you can see NSL dataset accuracy



In above screen CNN got 54% accuracy and then scroll down above text area to get BAT-MC accuracy



In above screen BAT-MC got 95% accuracy, Below code screen showing how we are creating BLSTM model to build attack detection model



In above screen read red colour comments to understand development of BAT-MC model

CONCLUSION

The current deep learning methods in the network traffic classification research don't make full use of the network traffic structured information. Drawing on the application methods of deep learning in the field of natural language processing, we propose a novel model BAT-MC via the two phase's learning of BLSTM and attention on the time series features for intrusion detection using NSL-KDD dataset. BLSTM layer which connects the forward LSTM and the backward LSTM is used to extract features on the the traffic bytes of each packet. Each data packet can produce a packet vector. These packet vectors are arranged to form a network flow vector. Attention layer is used to perform feature learning on the network flow vector composed of packet vectors. The above feature learning process is automatically completed by deep neural network without any feature engineering technology.This model effectively avoids the problem of manual design features. Performance of the BAT-MC method is tested by KDDTest+ and KDDTest-21 dataset. Experimental results on the NSL-KDD dataset indicate that the BAT-MC model achieves pretty

high accuracy. By comparing with some standard classifier, these comparisons show that BAT-MC models results are very promising when compared to other current deep learning-based methods. Hence, we believe that the proposed method is a powerful tool for the intrusion detection problem.

REFERANCES

1. B. B. Zarpelo, R. S Miani, C. T. Kawakani, and S. C. de Alvarenga, ‘‘A survey of intrusion detection in Internet of Things,’’ J. Netw. Comput. Appl., vol. 84, pp. 25–37, Apr. 2017.

2. B. Mukherjee, L. T. Heberlein, and K. N. Levitt, ‘‘Network intrusion detection,’’ IEEE Netw., vol. 8, no. 3, pp. 26–41, May 1994.

3. S. Kishorwagh, V. K. Pachghare, and S. R. Kolhe, ‘‘Survey on intrusion detection system using machine learning techniques,’’ Int. J. Control Automat., vol. 78, no. 16, pp. 30–37, Sep. 2013.

4. N. Sultana, N. Chilamkurti, W. Peng, and R. Alhadad, ‘‘Survey on SDN based network intrusion detection system using machine learning approaches,’’ Peer-to-Peer Netw. Appl., vol. 12, no. 2, pp. 493–501, Mar. 2019.

5. M. Panda, A. Abraham, S. Das, and M. R. Patra, ‘‘Network intrusion detection system: A machine learning approach,’’ Intell. Decis. Technol., vol. 5, no. 4, pp. 347–356, 2011.

6 W. Li, P. Yi, Y. Wu, L. Pan, and J. Li, ‘‘A new intrusion detection system based on KNN classification algorithm in wireless sensor network,’’ J. Electr. Comput. Eng., vol. 2014, pp. 1–8, Jun. 2014.

7.S. Garg and S. Batra, ‘‘A novel ensembled technique for anomaly detection,’’ Int. J. Commun. Syst., vol. 30, no. 11, p. e3248, Jul. 2017.

8. F. Kuang, W. Xu, and S. Zhang, ‘‘A novel hybrid KPCA and SVM with GA model for intrusion detection,’’ Appl. Soft Comput.., vol. 18, pp. 178–184, May 2014.

9. W. Wang, M. Zhu, X. Zeng, X. Ye, and Y. Sheng, ‘‘Malware traffic classification using convolutional neural network for representation learning,’’ in Proc. Int. Conf. Inf. Netw. (ICOIN), 2017, pp. 712–717.

10. P. Torres, C. Catania, S. Garcia, and C. G. Garino, ‘‘An analysis of Recurrent Neural Networks for Botnet detection behavior,’’ in Proc. IEEE Biennial Congr. Argentina (ARGENCON), Jun. 2016, pp. 1–6.