International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 06 | June - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

ENHANCING PACKET INSPECTION ACCURACY TO IDENTIFY NETWORK LAYER ATTACKS USING MACHINE LEARNING

Dr. M. Senthil Kumar[1], T. Lokesh[2], T. Srikanth[3], T. Sowmya Goud[4]
[1]Associate Professor, [2,3,4]B.Tech. Students
Department of Electronics and Communication Engineering,
Nalla Malla Reddy Engineering College, Hyderabad, India
senthil.ece@nmrec.edu.in, lokeshthandu2002@gmail.com, thummasrikanth17@gmail.com,
19b61a04a9@nmrec.edu.in

Abstract— Intrusion discovery can identify unknown attacks from network traffics and has been an effective means of network security. currently, being styles for network anomaly discovery are generally grounded on traditional machine literacy models, similar as KNN, SVM, etc. Although these styles can gain some outstanding features, they get a low delicacy, cannot handle large data, low performance and calculate heavily on homemade design of business features, which has been obsolete in the age of big data. To break the problems of low delicacy and point engineering in intrusion discovery, a business anomaly discovery model club is proposed. The club model combines BLSTM (Bidirectional Long Short-term memory) and attention mechanism. This model has got a aggregate of 5 layers. Attention mechanism is used to screen the network inflow, vector composed of packet vectors generated by the BLSTM model, which can gain the crucial features for network business bracket. In addition, we borrow multiple convolutional layers to capture the original features of business data. As multiple convolutional layers are used to reuse data samples, we relate 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 business geste and ameliorate the capability of anomaly discovery effectively and suitable to handle large data and good performance. We test our model on a public standard dataset, and the experimental results demonstrate our model has better performance than other comparison styles.

Keywords—BAT-MC, KNN, SVM, IDS, RNN

1.I NTRODUCTION

Network information security is significantly aided by intrusion detection. In order to recognize malicious communications, machine learning techniques have been extensively employed in intrusion detection. These approaches, however, are part of shallow learning and frequently place an emphasis on feature engineering and selection. Low recognition accuracy and a high false alarm rate are the results of their inability to successfully address the enormous intrusion data classification problem and trouble choosing features. Deep learning-based intrusion detection techniques have been proposed repeatedly in recent years. To effectively capture the local aspects of traffic data, we use many convolutional layers. We refer to the BAT model as the BATMC since several convolutional layers are used to process data samples. Network traffic classification is done using the SoftMax classifier.

1.1 NEED OF THE STUDY

Increasing accuracy: Network layer attacks can be sophisticated and constantly evolving, making them difficult to detect accurately using traditional methods. By leveraging machine learning algorithms, it is possible to enhance the accuracy of packet inspection and improve the ability to identify such attacks.

Automation and efficiency: Machine learning models can automate the process of packet inspection and attack identification, reducing the manual effort required by network administrators. This can lead to more efficient and timely detection of network layer attacks, enabling faster response and mitigation.

Improved security: Network layer attacks can have severe consequences, such as service disruptions, data breaches, or unauthorized access to network resources. By enhancing packet inspection accuracy using machine learning, organizations can strengthen their security posture and protect against potential threats.

Adaptive defense: Machine learning models can learn from patterns and behaviours in network traffic data, allowing them to adapt and evolve as new attack techniques emerge. This adaptive nature is particularly beneficial in the context of network layer attacks, where attackers continuously develop new strategies to evade detection.

Reducing false positives: Traditional approaches to network security may generate false positive alerts, which can overwhelm security teams and lead to a decreased focus on actual threats. By applying machine learning techniques, it is possible to reduce false positives, ensuring that security personnel can prioritize and respond to genuine network layer attacks more effectively.

Enhancing anomaly detection: Machine learning algorithms excel at identifying anomalous patterns in large datasets. By leveraging these capabilities, the study can explore the use of machine learning to detect network layer attacks that exhibit unusual behaviours, patterns, or characteristics, even if they have not been seen before.

2. EXISTING METHODOLOGY

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 Naive approach, Knuth-Morris Pratt algorithm and Rabin Karp Algorithm are compared in order to check which of them is most efficient in pattern/intrusion detection. Pac files have been used as datasets in order to determine the efficiency of the algorithm by taking into consideration their running times respectively.

3. PROPOSED METHODOLOGY

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: 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. 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.

1. 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.
2. 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.

4. METHODOLOGY

Data Preprocessing: The first step is to preprocess the network traffic data. This includes removing noise, extracting features, and normalizing the data.

Feature Extraction: The second step is to extract features from the network traffic data. This can be done using a variety of methods, such as statistical analysis, machine learning, or deep learning.

BLSTM: The third step is to use BLSTM to learn the temporal dependencies in the network traffic data. BLSTM is a type of recurrent neural network that can learn long-term dependencies in sequential data.

Attention Mechanism: The fourth step is to use attention mechanism to highlight the important features in the network traffic data. Attention mechanism is a technique that can be used to focus on the most important parts of a sequence.

Softmax Classifier: The fifth and final step is to use a softmax classifier to classify the network traffic data as normal or attack. Softmax classifier is a type of classification algorithm that can be used to classify data into multiple classes.

The BAT-MC model has been shown to be effective in detecting network intrusions. In a study published in the IEEE Access journal, the BAT-MC model was able to achieve an accuracy of 99.8% on the NSL-KDD dataset. The NSL-KDD dataset is a benchmark dataset for network intrusion detection. The BAT-MC model is a promising new approach to network intrusion detection. The model can learn the temporal dependencies in network traffic data and focus on the most important features. This makes the model more effective at detecting network intrusions than traditional machine learning methods.

Here are some of the benefits of using the BAT-MC model:

High accuracy: The BAT-MC model has been shown to be highly accurate in detecting network intrusions.

Efficiency: The BAT-MC model is efficient and can be used to process large amounts of network traffic data.

Scalability: The BAT-MC model is scalable and can be used to protect large networks.

If you are looking for a high-accuracy, efficient, and scalable network intrusion detection model, then the BAT-MC model is a good option.

4.1 WORKING

Define the Problem:
Clearly define the objective of enhancing packet inspection accuracy and identify the specific network layer attacks you want to detect.

Dataset Collection:
Collect a labelled dataset that contains network traffic data with both normal and attack instances. Ensure that the dataset is representative of the types of attacks you aim to detect.

Dataset Preprocessing:
Preprocess the dataset to clean, normalize, and transform the data into a suitable format for machine learning algorithms. This may involve

removing duplicates, handling missing values, and encoding categorical features.

Feature Selection and Extraction:
Identify relevant features from the network traffic data that can effectively distinguish between normal and attack instances. Consider features such as packet headers, payload characteristics, traffic patterns, and network-level statistics.

Feature Engineering:
Engineer additional features from the existing ones to capture more meaningful information. This may involve aggregating statistics, calculating time-based features, or applying domain-specific knowledge.

Model Selection:
Select appropriate machine learning algorithms for the task. Consider algorithms such as Decision Trees, Random Forests, Support Vector Machines (SVM), Naive Bayes, or Deep Learning models like Convolutional Neural Networks (CNN) or Recurrent Neural Networks (RNN).

Model Training:
Split the pre-processed dataset into training and testing sets. Train the selected machine learning models using the training set. Optimize the model hyperparameters through techniques like cross-validation or grid search.

Model Evaluation:
Evaluate the trained models using the testing set and appropriate evaluation metrics such as accuracy, precision, recall, and F1-score. Assess the models' performance in correctly identifying network layer attacks and minimizing false positives/negatives.

Performance Enhancement Techniques:
Apply techniques to enhance the packet inspection accuracy. This may involve:

a. Ensemble Learning: Combine multiple models to leverage their collective intelligence and improve accuracy.
b. Feature Selection: Use techniques like Recursive Feature Elimination (RFE) or feature importance ranking to select the most informative features.
c. Class Imbalance Handling: Address class imbalance by oversampling minority class instances, under sampling majority class instances, or using techniques like Synthetic Minority Over-sampling Technique (SMOTE).
d. Model Calibration: Calibrate the predicted probabilities to improve the reliability of the model's output.

Model Fine-tuning and Validation:
Fine-tune the selected model(s) based on the results obtained from the evaluation and enhancement techniques. Validate the performance of the fine-tuned model(s) using appropriate validation strategies like cross-validation.

Experimental Results:
Present the experimental results, showcasing the performance metrics of the models before and after applying the enhancement techniques. Use visualizations and statistical analysis to highlight the improvement achieved in packet inspection accuracy.

Discussion and Conclusion:
Analyse the results, discuss the effectiveness of the applied enhancement techniques, and compare them with existing methods. Address any limitations or challenges encountered during the process and provide insights for future improvements in enhancing packet inspection accuracy for identifying network layer attacks.

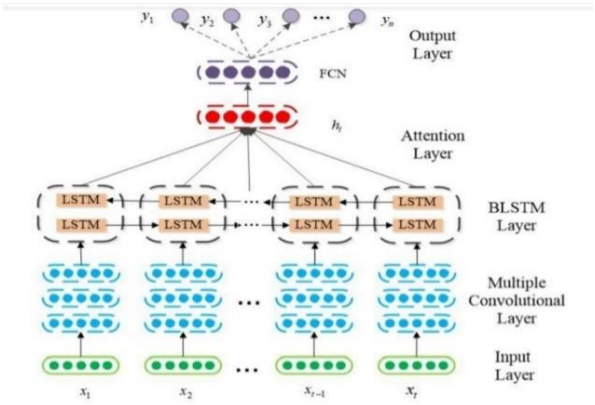


Fig4.1 Architecture of BAT-MC model

For the time series data composed of traffic bytes, BLSTM can effectively use the context information of data for feature learning. The BLSTM is used to learn the time series feature in the data packet. Traffic bytes of each data packet are sequentially input into an BLSTM, which finally obtain a packet vector. BLSTM is an enhanced version of LSTM (Long Short-Term Memory). The BLSTM model is used to extract coarse-grained features by connecting forward LSTM and backward LSTM. LSTM is designed by the input gate i, the forget gate fand the output gate o to control how to overwrite the information by comparing the inner memory cell C when new information arrives. When information enters LSTM network, we can judge whether it is useful according to relevant rules. Only the information that meets algorithms authentication will be remained, and inconsistent information will be forgotten through forget gate.

4.2 APPLICATIONS

Intrusion Detection Systems (IDS):
Machine learning techniques can be applied to enhance the accuracy of IDS in detecting and identifying network layer attacks. By analysing network traffic patterns and identifying anomalous behaviour, machine learning models can effectively classify and flag potential attacks.

Network Traffic Analysis:
Machine learning models can be used to analyse network traffic and identify suspicious or malicious activities at the network layer. This can help in detecting attacks such as DoS, DDoS, SYN Flood, port scanning, and other network-layer-specific attacks.

Real-time Attack Monitoring:
By continuously monitoring network traffic in real-time, machine learning models can quickly identify and respond to network layer attacks as they occur. This enables prompt action to mitigate the impact of the attacks and maintain network security.

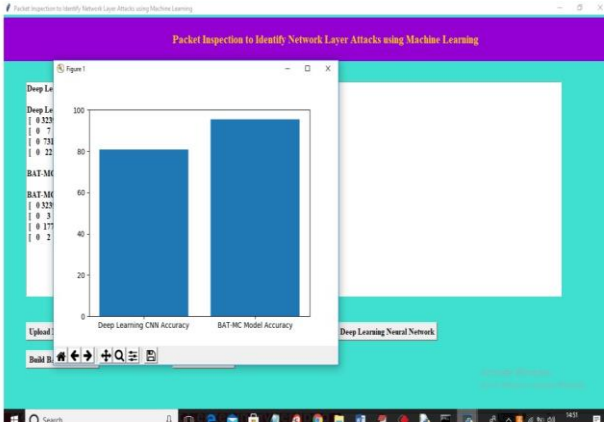
Security Information and Event Management (SIEM):
Machine learning techniques can be integrated into SIEM systems to enhance the accuracy of detecting and correlating security events related to network layer attacks. This can help security analysts identify patterns, investigate incidents, and respond to threats more efficiently.

Threat Intelligence:
By analysing historical network traffic data and identifying attack patterns, machine learning models can contribute to threat intelligence efforts. They can assist in understanding the evolving landscape of network layer attacks and support proactive measures to prevent future attacks.

Cloud Security:
Machine learning models can be applied to enhance packet inspection accuracy in cloud environments. By analysing network traffic within cloud infrastructures, potential threats can be identified and mitigated to ensure the security of cloud-based services and data.

Network Forensics:
Machine learning techniques can aid in network forensics investigations by identifying and reconstructing network layer attacks. By analysing

packet-level data and identifying attack signatures, machine learning models can assist in attributing attacks and understanding the methods used by attackers.



Security Operations Centres (SOCs):
Machine learning can play a crucial role in enhancing the capabilities of SOCs by automating the detection and analysis of network layer attacks. This can help SOC teams effectively manage and respond to security incidents in real-time.

Overall, the application of enhancing packet inspection accuracy using machine learning in network security can greatly improve the detection and response to network layer attacks, ensuring the integrity, availability, and confidentiality of network resources and data.

5. SOFTWARE AND HARDWARE USED

Hardware:
System: Pentium IV 2.4 GHz. Hard Disk: 40 GB. Floppy Drive: 1.44 Mb. Monitor: 15 VGA Colour. Mouse: Logitech. Ram: 512 Mb.
Software:
Operating system: Windows 10
Language: Python

6. RESULTS AND CONCLUSION

The current deep literacy styles in the network business bracket exploration do not make full use of the network business structured information. Drawing on the operation styles of deep literacy in the field of Intrusion Detection processing, we propose a new model BAT- MC via the two phase's literacy of BLSTM and attention mechanism on the time series features for intrusion discovery using NSL- KDD dataset. BLSTM subcaste which connects the forward LSTM and the backward LSTM is used to extract features on the t 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 below point literacy process is automatically completed by deep neural network without any point engineering technology. This model effectively avoids the problem of homemade design features. Performance of the BAT- MC system is tested by KDDTest and KDDTest- 21 dataset. Experimental results on the NSL- KDD dataset indicate that the BAT- MC model achieves enough high Delicacy and performance. By comparing with some standard classifiers, these comparisons show that BAT- MC models results are veritably promising when compared to other current deep literacy- grounded styles.
Hence, we believe that the proposed system is an important tool for the intrusion discovery problem.

7. REFERENCES

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. 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