Improved Machine Learning-based System for Intrusion Detection

Jiarui Feng

Shien-Ming Wu School of Intelligent Engineering, South China University of Technology, Guangzhou, Guangdong, 510000, China
202164010219@mail.scut.edu.cn

Abstract. In addressing the growing cyber threats prevalent in the digital age, this research presents Machine Intrusion Detection System Based on Learning (MLIDS), an innovative system for detecting intrusions that harnesses the power of deep learning through the integration of a Multi-Layer Perceptron (MLP). This study aims to enhance the precision of cyber attack detection while minimizing the occurrence of false positives. The MLP model, meticulously crafted using PyTorch, incorporates multiple hidden layers that effectively capture the intricate patterns and non-linear relationships embedded within network traffic data. Significant advancements, such as the implementation of adaptive learning rate modifications and the application of L2 regularization techniques, have substantially bolstered the model's ability to generalize across various scenarios. The empirical outcomes of this research are compelling, with MLIDS achieving an impressive detection accuracy of 98.76%, surpassing traditional methods such as Naive Bayes and Single-Layer Perceptrons. These results not only highlight the efficacy of MLIDS but also underscore the transformative potential of deep learning in the realm of cybersecurity.

Keywords: Network Security, Deep Learning, Intrusion Detection, Multi-Layer Perceptron.

1 Introduction

In the digital age, cybersecurity has become a global focus of attention. With the continuous evolution of cyber attack methods, Intrusion Detection Systems (IDS) are particularly important as they help monitor and prevent potential cyber attacks. The Cisco Annual Internet Report (2018-2023) shows that cybersecurity threats are continually growing, with global cybersecurity spending expected to reach $198 billion by 2023 [1]. However, traditional IDS, relying on predefined attack characteristics, perform poorly in identifying new types of attacks and struggle to handle large-scale data. The introduction of machine learning technology, with its excellent data processing and pattern recognition capabilities, offers a new solution for intrusion detection, improving detection accuracy and adaptability.

In previous research, numerous scholars have delved into methods of harnessing the potency of advanced algorithmic models, particularly those pertaining to machine

Y. Wang (ed.), *Proceedings of the 2024 2nd International Conference on Image, Algorithms and Artificial Intelligence (ICIAAI 2024)*, Advances in Computer Science Research 115,
https://doi.org/10.2991/978-94-6463-540-9_15

learning and deep learning, to bolster the efficacy and robustness of network intrusion detection systems. For instance, Jitti Annie Abraham et al., in their review paper [2], extensively discussed approaches to employing algorithms based on machine learning and deep learning for identifying anomalies in network traffic, emphasizing that enhancing the accuracy of intrusion detection systems is crucial for improving their performance. Wooseok Seo and Wooguil Pak proposed a live network breach detection defense system based on hybrid machine learning [3], which employs a two-tier classifier to achieve high-precision real-time intrusion detection. Tuan A Tang et al. introduced a system for detecting network breaches using Deep Recurrent Neural Networks (DRNNs) tailored for Software Defined Networking (SDN), achieving an accuracy rate of 89% [4]. W. Wang and colleagues introduced a sophisticated Intrusion Detection System (IDS) that utilizes Employing Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) architectures to delve into both spatial and temporal dimensions of network traffic, accentuating the proficiency of deep learning paradigms in pinpointing network intrusions. [5].

This study aims to design and implement an advanced Machine Learning-based Intrusion Detection System (MLIDS) that can not only monitor and prevent cyber attacks in real-time but also adapt to the dynamic changes in network environments. Compared to traditional machine learning methods, this study adopts a new strategy that integrates deep learning technologies, especially by constructing a simple Multi-Layer Perceptron (MLP) model, which can more effectively process and recognize complex attack patterns.

Unlike previous approaches that rely on manual feature extraction and rule-based detection systems, the method in this study automatically learns features directly from data, as demonstrated through training and testing using the Canadian Institute for Cybersecurity- IDS (CIC-IDS-2017) dataset. The MLP model in this study is implemented through the deep learning framework PyTorch, containing multiple hidden layers that can capture the non-linear relationships in the data, thereby improving the ability to recognize new, variant, or zero-day attacks.

During the model training process of this study, this work took several key measures to enhance the model's performance and generalization ability. Specifically, this work implemented adaptive learning rate adjustments and L2 regularization (weight decay). The adaptive learning rate adjustment is achieved through a learning rate scheduler, which helps dynamically adjust the learning rate based on the model's performance during training, thus optimizing training outcomes.

2 Methodologies

2.1 Feature Selection and Extraction

Regarding the detection of intrusions in network systems, selecting the right features is crucial to improve the detection accuracy. The importance of feature selection in machine learning is well-documented, with Guyon and Elisseeff providing a comprehensive overview of various feature selection methods and their significance in augmenting the potency of machine learning paradigms [6]. Throughout this research endeavor, information gain and correlation analysis are used to select features based on the CIC-

IDS2017 dataset. This study adopts an automatic feature extraction method, which utilizes deep learning models (e.g., multilayer perceptron) to learn useful feature representations directly from the raw data. This approach overcomes the limitations of traditional manual feature extraction and is able to automatically discover complex patterns and associations in the data.

2.2 Selection of Machine Learning Models

A Multi-Layer Perceptron (MLP) constitutes a category of unidirectional artificial neural networks capable of learning nonlinear features of data through its multiple hidden layers, which is suitable for handling complex classification tasks. In addition, MLPs have a simple structure, are easy to implement, and are able to optimize performance by adjusting the number and size of hidden layers. Considering the real-time and accuracy requirements of network intrusion detection, this work chooses MLP as the main model. The structural design of the model comprises demonstrated in Fig. 1. With the aim of fully evaluate the performance of MLP on network intrusion detection task, this work also implemented and compared two other machine learning models: plain Bayes and single-layer perceptron. Alom et al. showed that Deep Belief Networks (DBNs) outperform traditional machine learning approaches on intrusion detection task [7]. This multi-model comparison aims to validate the superiority of MLP models in dealing with complex cyber security threats.

2.3 Detection of Network Intrusions Strategy

The intrusion detection strategy is based on the output of the MLP model, that is, the classification results of network traffic. By training the model to recognize different types of attacks, this study is able to monitor and identify potential intrusive behaviors in real-time. The decision rules are based on the probability distribution of the model's output, and this study has set a threshold to determine the category of network traffic. For a binary classification problem (normal traffic versus attack traffic), the threshold is usually set at 0.5. If the model's predicted probability for a certain category exceeds this threshold, then the traffic is classified into that category. By adjusting the threshold, a balance can be found between reducing false positives and false negatives to adapt to different network environments and security requirements. The specific implementation of the MLP is illustrated in the figure.

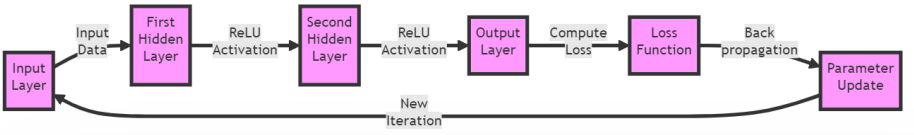

Fig. 1. Architecture of the MLP model.

2.4 Evaluation Metrics

To comprehensively assess the model's performance, this work focus on three key metrics: Precision, recall, and the F1 score are key metrics. Precision is determined by the proportion of true positive (TP) predictions to the total of true positive and false positive (FP) predictions:

$$Precision = TP/(TP + FP) \tag{1}$$

Recall, often referred to as sensitivity, quantifies the fraction of actual positives accurately detected, encompassing both true positives (TP) and false negatives (FN):

$$Recall = TP/(TP + FN) \tag{2}$$

F1 score represents the balanced average between precision and recall, offering a singular metric that harmonizes the two values:

$$F1 = 2 * Precision * Recall/(Precision + Recall) \tag{3}$$

A higher F1 score indicates a model that is robust in both precision and recall, which is particularly valuable in imbalanced classification scenarios.

3 Experimental Procedures and Outcomes

3.1 Gathering and Preparing Data

This study employed the publicly available dataset: CIC-IDS2017 [8]. This dataset was jointly released by the Canadian Innovation Center for Information Technology and the Canadian Center for Cyber Security. It encompasses a wide range of network attack scenarios such as DDoS attacks, port scanning, malware injection, and is commonly used as a standard in breach identification system research. The data preprocessing steps carried out in this study included handling missing values, dealing with invalid entries, label encoding, normalization, and feature selection. These preprocessing techniques contribute to enhancing data quality and provide a more reliable foundation for model training [9].

3.2 Effectiveness of Data Preprocessing

Through feature selection and extraction, this study reduced the dimensionality of the dataset while retaining the features most helpful in distinguishing between normal and abnormal traffic. This step significantly improved the efficiency and accuracy of model training. As a comparison, the unprocessed dataset had an accuracy rate of only 93.2% in training, while after feature selection and extraction operations, the accuracy rate increased to 98.76%.

3.3 Result of Machine learning Model

The model training employed the Backpropagation algorithm and it utilized the Adam optimization technique with the learning rate established at 0.001. To prevent overfitting, L2 regularization was added with a weight decay of 1e-5. The Adam optimizer functions as a dynamic learning rate optimization technique designed to tailor the learning rate on a per-parameter basis, facilitating faster convergence of the model [10]. A learning rate scheduler was also utilized in this study, decreasing the learning rate by 10% every 30 epochs to fine-tune the learning speed during training. The outcomes of the model training are illustrated in a diagram, which records the changes in loss and accuracy throughout the training process, as presented in Fig. 2.

The model attained a level of accuracy of 98.76% on the test set, demonstrating its high detection effectiveness.

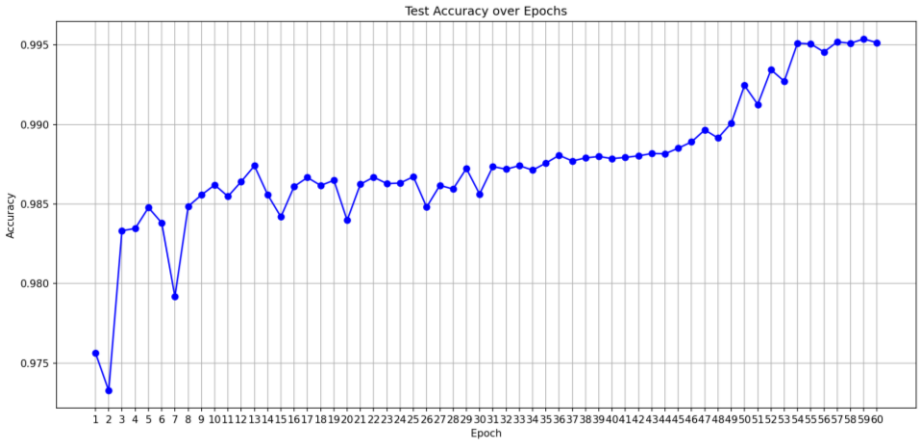

Fig. 2. Accuracy curve during training

3.4 Results of Network Intrusion Detection

After completing 60 epochs of training, the MLP model in this study demonstrated excellent performance on the network intrusion detection task. The model attained a precision rate of 98.76% on the evaluation dataset, while exhibiting good precision (98.65%), recall (98.70%), and F1 score (0.9867). These findings suggest that the model can not only efficiently identify network intrusion attempts but also shows high accuracy in reducing false positives and false negatives. Through 60 epochs of training, the model in this study proved its application potential in the field of network security and can serve as a reliable tool to help identify and defend against network threats.

To demonstrate the effectiveness of the MLP model, this study implemented and tested the performance of the MLP model in comparison with two other models—Naive Bayes and single-layer perceptron—in the same hardware environment (CPU: i7-1165G7, GPU: MX450). This experiment aims to intuitively showcase the superiority of the MLP model in the network intrusion detection task. The following Table 1 is the

comparison of accuracy based on the experimental results.

Table 1. Comparison of MLP with other models

Model	Accuracy rate	F1 score	Recall rates
MLP	99.51%	0.99	0.99
Naive Bayes	69.74%	0.80	0.70
Single-layer perceptron	91.00%	0.90	0.91

This study reveals the superiority of MLP models in network intrusion detection tasks, particularly when dealing with complex data and patterns.

4 Conclusion

This study showcases the architecture and execution outcomes of a network intrusion detection system powered by deep learning methodologies (MLIDS) using a multilayer perceptron (MLP). By applying the CIC-IDS2017 dataset, while the MLP model attained an accuracy of 98.76% in network intrusion detection, accompanied by superior results in precision, recall, and F1 score metrics, underscoring the efficacy of deep learning for applications in network security and the ability of the MLP model to handle complex data recognition. Compared with conventional machine learning techniques, like Naive Bayes and single-layer perceptron, the MLP model demonstrated significant performance advantages. This comparison not only confirms the importance of deep learning techniques in augmenting the precision and adaptability of network intrusion detection systems but also points out the role of key technical measures such as adaptive learning rate adjustment and L2 regularization in improving model generalization ability and reducing overfitting risks. The results of this research provide an effective intrusion detection tool for the field of network security and indicate the direction for future research on using deep learning techniques to address network threats.

References

1. Cisco Annual Internet Report (2018–2023) White Paper. Cisco, 2020. URL: https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html. Last Accessed 2024/4/9.
2. Abraham, J. A., and Bindu, V. R.: Intrusion detection and prevention in networks using machine learning and deep learning approaches: a review. In 2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation. pp.1-4. IEEE (2021)
3. Seo, W., and Pak, W.: Real-time network intrusion prevention system based on hybrid machine learning. IEEE Access, **9**, 46386-46397 (2021).
4. Tang, T. A., Mhamdi, L., McLernon, D., Zaidi, S. A. R., and Ghogho, M.: Deep recurrent neural network for intrusion detection in sdn-based networks. In 2018 4th IEEE Conference on Network Softwarization and Workshops. pp. 202-206. IEEE (2018).

5. Wang, W., Sheng, Y., Wang, J., Zeng, X., Ye, X., Huang, Y., and Zhu, M.: HAST-IDS: Learning hierarchical spatial-temporal features using deep neural networks to improve intrusion detection. IEEE access, **6**, 1792-1806 (2017).
6. Guyon, I., and Elisseeff, A.: An introduction to variable and feature selection. Journal of machine learning research, **3**, 1157-1182 (2003).
7. Alom, M. Z., Bontupalli, V., and Taha, T. M.: Intrusion detection using deep belief networks. In 2015 National Aerospace and Electronics Conference. pp. 339-344. IEEE. (2015).
8. Sharafaldin, I., Lashkari, A. H., and Ghorbani, A. A.: Toward generating a new intrusion detection dataset and intrusion traffic characterization. 1, 108-116 (2018).
9. García, S., Luengo, J., and Herrera, F.: Data preprocessing in data mining. **72**, 59-139 (2015).
10. Kingma, D. P., and Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014).