Enhancing Intrusion Detection Systems Using Hybrid Deep Learning Algorithms

C.Vejhayac Khumar,

Teaching Assistant,Surya Engineering College,Erode.

Abstract: This research focus Traditional intrusion detection systems (IDS) frequently struggle to stay accurate and flexible as cyber threats grow more complex. In order to improve intrusion detection capabilities, this study presents a hybrid deep learning technique that combines Convolutional Neural Networks (CNN) with Long Short-Term Memory (LSTM) networks. The model is trained and evaluated on popular datasets such as CICIDS2017 and NSL-KDD. Metrics like accuracy, precision, recall, and F1-score are used to assess its performance. The findings demonstrate that, in comparison to standalone models, this method produces higher detection rates and fewer false positives. This hybrid architecture offers a robust and scalable response to the current network security issues by utilizing both spatial and temporal data features.

Keywords: Intrusion Detection System (IDS),Hybrid Deep Learning, CNN-LSTM, Network Security, Cyber Threat Detection

INTRODUCTION

This research aims to in our fast-paced digital world, the explosion of internet technologies and interconnected systems has really widened the playing field for cybercriminals. With the rise of sophisticated and ever-evolving threats—like zero-day attacks, advanced persistent threats (APTs), ransom ware, and phishing—the demand for strong and flexible cyber security measures is more urgent than ever. One of the key players in this arena is the Intrusion Detection System (IDS), which is crucial for keeping an eye on and analyzing network traffic to spot any suspicious behavior that might suggest unauthorized access or malicious activity. Traditional IDS methods generally fall into two categories: signature-based and anomaly-based. Signature-based IDS depend on known patterns or attack signatures to pinpoint threats.

While do a good job against familiar attacks, they struggle when faced with new or disguised threats. On the flip side, anomaly-based IDS try to model what normal network behavior looks like and flag anything that deviates from that. They can catch previously unseen threats, these systems often deal with a high rate of false positives and need regular retraining to keep up with the ever-changing network landscape. To overcome the shortcomings of traditional IDS, there's been a growing interest in using machine learning (ML) and deep learning (DL) techniques to boost intrusion detection capabilities.

These methods can learn intricate patterns from vast amounts of data and adapt better to unknown attacks. Among these techniques, deep learning methods like Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks have shown great potential because they can capture both

spatial and temporal dependencies effectively. CNNs are particularly good at picking up local spatial features and have been traditionally used in image processing tasks. The knack for identifying localized patterns and correlations makes them a great fit for analyzing structured network data, such as traffic flow, packet headers, and other feature-rich representations. CNNs can effectively spot abnormal patterns in this data, enhancing our ability to detect potential threats.

Research Question:

1. How does the CNN-LSTM hybrid model improve IDS performance?
2. What is the benefit of combining spatial and temporal features?
3. How well does the model perform on NSL-KDD and CICIDS2017 datasets?

LITERATURE REVIEW:

Intrusion Detection Systems (IDS) play a crucial role in the world of cybersecurity, focusing on spotting unauthorized access and unusual network activities. Traditional methods, like signature-based and anomaly-based approaches, have laid the groundwork for protecting networks. However, these systems often struggle to identify new types of attacks and tend to produce a high number of false positives, which can hinder their effectiveness in ever-changing environments (Garcia-Teodoro et al., 2009). With the rise of artificial intelligence, machine learning (ML) techniques such as Decision Trees, Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Naive Bayes are being increasingly utilized in IDS. These models have the edge over static methods because they learn patterns from data. Still, they depend heavily on feature engineering and might miss out on capturing the complex, nonlinear patterns that exist in network traffic (Sangkatsanee et al., 2011).

The advent of deep learning (DL) presents a powerful alternative, thanks to its ability to automatically extract high-level features from large datasets. Convolutional Neural Networks (CNN), originally designed for image processing, have shown great promise in recognizing spatial patterns within network data. CNNs have been effectively used in intrusion detection to classify traffic as either benign or malicious by converting traffic features into 2D representations (Kim et al., 2016).

On the other hand, Recurrent Neural Networks (RNN), especially Long Short-Term Memory (LSTM) networks, excel at modeling sequential and temporal data. LSTM-based IDS models have demonstrated their effectiveness in detecting time-dependent attack patterns, such as port scanning and denial-of-service (DoS) attacks (Yin et al., 2017). However, relying solely on LSTM models may not adequately capture spatial dependencies, which can limit their performance. To tackle these gaps, recent research has delved into hybrid deep learning models that blend CNN and LSTM networks, effectively harnessing both spatial and temporal features. For example, Vinayakumar et al. (2019) created a CNN-LSTM hybrid model that achieved

impressive detection accuracy on benchmark datasets. In a similar vein, Alom et al. (2018) introduced a deep learning architecture that merges CNN and RNN for cybersecurity applications, showcasing enhanced performance compared to standalone models.

When it comes to evaluating IDS performance, benchmark datasets like NSL-KDD and CICIDS2017 are frequently utilized. NSL-KDD is an upgraded version of KDD'99, addressing issues of redundancy and imbalance, while CICIDS2017 captures modern attack scenarios and realistic traffic patterns. These datasets have become the go-to standards for comparing IDS models in academic research.

SOFTWARE AND HARDWARE CONFIGURATION

The TensorFlow 2.8 framework along with its user-friendly Keras API to efficiently design, train, and evaluate our deep learning architecture. We relied on essential libraries like NumPy and Pandas for dataset manipulation and preprocessing, while Scikit-learn helped us with data splitting, encoding categorical variables, normalization, and calculating evaluation metrics. To visualize the training progress and results, we used Matplotlib and Seaborn. All experiments were carried out on a high-performance workstation equipped with

1. CPU: Intel Core i7-9700K running at 3.6 GHz (8 cores, 8 threads)
2. GPU: NVIDIA GeForce RTX 2080 Ti with 11 GB of VRAM,
3. RAM: 32 GB DDR4, more than enough to manage large datasets in memory during training sessions.
4. Storage: 1 TB SSD for quick data loading and saving model checkpoints. Operating System: Windows 10 Pro 64-bit, ensuring reliable support for all necessary drivers and libraries.

The datasets were divided into training, validation, and testing sets in an 80:10:10 ratio. The hybrid model architecture utilized convolutional layers to pull out spatial features from network traffic data, along with LSTM layers to capture temporal dependencies. This was followed by fully connected layers for the final classification. The model was trained using the Adam optimizer, with categorical cross-entropy serving as the loss function. A batch size of 64 was selected to strike a balance between computational efficiency and memory limitations, and the training process was carried out over a maximum of 50 epochs, incorporating early stopping based on validation loss to avoid overfitting. Additionally, dropout layers were included for regularization. To assess performance, we used metrics like accuracy, precision, recall, and F1-score, which together offer a comprehensive evaluation of the model's intrusion detection capabilities, ensuring high detection rates while keeping false positives to a minimum.

INTRUSION DETECTION DATASETS

The hybrid CNN-LSTM model was put to the test using two popular intrusion detection datasets: NSL-KDD and CICIDS2017. To gauge its performance, we looked at key metrics like accuracy, precision, recall, and F1-score, giving us a well-rounded view of how well it detects threats.

When we examined the NSL-KDD dataset, the hybrid model scored an impressive accuracy of 97.8%, surpassing the traditional standalone CNN and LSTM models, which managed accuracies of 94.5% and 95.3%, respectively. The improvements in precision and recall were notable too, as the hybrid approach effectively minimized both false positives and false negatives compared to the individual models. This clearly shows the benefits of merging spatial and temporal feature extraction, enabling the model to better recognize complex attack patterns.

Turning to the CICIDS2017 dataset, which features a wider range of recent attack types, the model achieved an accuracy of 96.5%, with precision and recall values hitting 95.8% and 96.2%, respectively. These findings reinforce that the hybrid CNN-LSTM architecture consistently delivers strong detection performance across various datasets, highlighting its robustness and adaptability.

EXPERIMENTAL RESULTS AND INSIGHTS

The convolutional layers played a key role in efficiently extracting spatial features from network traffic data, while the LSTM layers adeptly captured temporal dependencies in sequential inputs, boosting the model's capability to identify both known and unknown attack patterns. Additionally, incorporating dropout and early stopping mechanisms helped to curb overfitting, ensuring the model performed well on new, unseen test data.

A comparative analysis with other cutting-edge models from recent studies shows that our proposed hybrid model significantly reduces false alarm rates while still maintaining high detection accuracy. This is particularly important in real-world cybersecurity scenarios, where an overload of false positives can burden analysts and undermine system trust.

CONCLUSION

In this research, we proposed and evaluated a hybrid deep learning model that combines Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks for intrusion detection in network security. As cyber threats become more complex and numerous, there's a growing need for advanced and adaptable detection systems that can accurately spot a wide variety of malicious activities while keeping false alarms to a minimum. Unfortunately, traditional intrusion detection systems (IDS) often struggle to maintain high accuracy and flexibility when confronted with sophisticated attacks and ever-

changing threat patterns. This study set out to tackle these issues by harnessing the unique strengths of both CNN and LSTM architectures within a single framework.

The CNN layers in our model did a great job of extracting spatial features from the input network traffic data, pinpointing local patterns that signal specific types of attacks. At the same time, the LSTM layers were able to capture the temporal dependencies and sequential characteristics of the traffic, enabling the model to recognize complex attack behaviors that develop over time. This hybrid approach provided a more nuanced understanding of both immediate and sequential attack signatures, which are often overlooked by standalone models.

Extensive experiments using two benchmark datasets, NSL-KDD and CICIDS2017, which cover a wide range of network traffic and attack scenarios. Our hybrid model consistently outshone the individual CNN and LSTM models across key performance metrics like accuracy, precision, recall, and F1-score. One of the standout features of the hybrid model was its impressive reduction in false positive rates, which is vital for real-world applications since too many false alarms can overwhelm security analysts and erode trust in the intrusion detection system (IDS). Additionally, the model proved to be highly adaptable, maintaining strong performance on both datasets despite their varying complexities and types of attacks.

Several technical choices played a key role in the model's success. We implemented preprocessing steps such as one-hot encoding for categorical features and Min-Max normalization for numerical data, which helped stabilize training and improve convergence. The Adam optimizer was used to ensure efficient learning with adaptive learning rates, while dropout regularization and early stopping were effective in preventing over fitting. The high-performance hardware setup, complete with GPU acceleration, allowed for the efficient training of this computationally intensive deep learning model. These encouraging results, there are some limitations that point to future research opportunities. While the model performed well on benchmark datasets, real-world network environments tend to be more dynamic and diverse. Future efforts could focus on deploying and testing the model in live network settings to assess its real-time detection capabilities and adaptability.

REFERENCE

1. Aljawarneh, S., Aldwairi, M., & Yassein, M. B. (2018). Anomaly-based intrusion detection system through feature selection analysis and building hybrid efficient model. *Journal of Computational Science, 25*, 152–160. https://doi.org/10.1016/j.jocs.2017.03.006
2. Yin, C., Zhu, Y., Fei, J., & He, X. (2017). A deep learning approach for intrusion detection using recurrent neural networks. *IEEE Access, 5*, 21954–21961. https://doi.org/10.1109/ACCESS.2017.2762418

3. Shone, N., Ngoc, T. N., Phai, V. D., & Shi, Q. (2018). A deep learning approach to network intrusion detection. *IEEE Transactions on Emerging Topics in Computational Intelligence, 2*(1), 41–50. https://doi.org/10.1109/TETCI.2017.2772792
4. Moustafa, N., & Slay, J. (2015). UNSW-NB15: A comprehensive data set for network intrusion detection systems. *2015 Military Communications and Information Systems Conference (MilCIS)*, 1–6. https://doi.org/10.1109/MilCIS.2015.7348942
5. Roy, S., Cheung, S., Sharma, S., & Saha, D. (2021). A CNN and LSTM based hybrid deep learning model for anomaly detection in network traffic. *Computer Communications, 175*, 104–113. https://doi.org/10.1016/j.comcom.2021.04.004
6. Dhanabal, L., & Shantharajah, S. P. (2015). A study on NSL-KDD dataset for intrusion detection system based on classification algorithms. *International Journal of Advanced Research in Computer and Communication Engineering, 4*(6), 446–452.
7. Sharafaldin, I., Lashkari, A. H., & Ghorbani, A. A. (2018). Toward generating a new intrusion detection dataset and intrusion traffic characterization. *Proceedings of the 4th International Conference on Information Systems Security and Privacy (ICISSP 2018)*, 108–116.