Smart Networks Traffic Analysis and Anomaly Detection using Bidirectional LSTM Networks with CUDA Kernel Utilization

Francisco Erialdo Domingos Freitas[1], **Antônio Wendell de Oliveira Rodrigues**[1]

[1]Graduate Program in Computer Science (PPGCC)
Federal Institute of Education, Science and Technology (IFCE)
Fortaleza – CE – Brazil

francisco.erialdo.domingos07@aluno.ifce.edu.br, wendell@ifce.edu.br

Abstract. *Modern networks, particularly those involving the Internet of Things (IoT), face significant challenges in traffic management, classification, and security. Efficient traffic classification and anomaly detection are critical for protecting these networks against malicious activities. In this work, we propose an approach based on Bidirectional Long Short-Term Memory (Bi-LSTM) networks, leveraging CUDA acceleration in PyTorch to optimize training performance and model accuracy. Our method captures contextual information from both past and future sequences, enhancing detection capabilities. Experimental results on an industrial IoT dataset demonstrate superior accuracy, recall, and F1 score compared to conventional LSTM models, highlighting the potential of the proposed solution for improving security and reliability in smart network environments.*

1. Introduction

1.1. General Context and Problem Characterization

The continuous advancement of communication technologies and the expansion of Internet of Things (IoT) networks have brought critical challenges related to security and traffic classification. Privacy breaches in these ecosystems generate distrust among users [Rizi and Seno 2022], making it essential to mitigate security risks to protect critical infrastructure, ensure urban resilience, and maintain optimal network performance [Rajasinghe et al. 2018]. In this context, network administrators and security professionals seek solutions capable of classifying traffic and identifying the cause of malicious behavior. This classification is fundamental for both security and compliance with Quality of Service (QoS) requirements, ensuring the low latency necessary for real-time applications (such as VoIP and IoT) at the expense of less sensitive traffic [Azab et al. 2024]. Thus, accurate data classification and anomaly detection become vital to differentiate benign from malicious traffic, allowing for real-time threat mitigation and optimization of bandwidth usage.

1.2. Related works

Recent works have explored various techniques for smart network anomaly detection, including Machine Learning and Deep Packet Inspection [Azab et al. 2024, Abbas et al. 2024]. Specific advancements in Recurrent Neural Networks, such as

sparse RNN implementations [Keirsbilck et al. 2019] and GPU Kernel optimizations [Gale et al. 2020], have improved computational efficiency. Furthermore, LSTM-based classifiers have shown promise in SDN-enabled IoT frameworks for identifying botnet attacks [Tayfour et al. 2023].

2. Methodology

To address traffic classification and anomaly detection in smart networks, this paper explores the application of Bidirectional Long Short-Term Memory (Bi-LSTM) networks implemented with CUDA support in PyTorch. Bi-LSTMs are highly suited for sequence-based data due to their ability to capture contextual information from both past and future states. By leveraging CUDA's computational power, we optimize the training process to achieve faster convergence and higher model accuracy, ultimately enhancing the security and robustness of smart network environments.

2.1. Bi-LSTM Overview

Bidirectional Long Short-Term Memory (Bi-LSTM) is an advanced Recurrent Neural Network architecture designed to learn long-term dependencies by processing input sequences in both chronological (forward) and reverse (backward) directions simultaneously [Zhao 2023]. For each step in a sequence of network traffic features, the forward layer computes the hidden state based on past context, while the backward layer captures future context. This dual temporal analysis successfully mitigates the gradient vanishing problem and enables a comprehensive extraction of protocol syntax and communication patterns [Yang et al. 2022].

The forward hidden state (h_t^f) in a Bidirectional Long Short-Term Memory (Bi-LSTM) network is computed by applying a nonlinear activation function, typically the hyperbolic tangent ($\tanh$), to a linear combination of the current input (x_t), the previous hidden state (h_{t-1}^f), and a bias term (b_h^f). The (W_{xh}^f) and (W_{hh}^f) are weight matrices associated with the input and the hidden state, respectively.

2.2. Materials and Methods

Figure 1 presents the general workflow adopted in this study. Initially, the input dataset, composed of IoT Modbus traffic records, undergoes a preprocessing phase where data cleaning and feature extraction are performed. Following preprocessing, the data is used to train the Bidirectional Long Short-Term Memory (Bi-LSTM) model, leveraging CUDA acceleration to optimize computational performance. After training, the model is evaluated using key performance metrics such as accuracy, recall, and F1 score. Finally, the results are analyzed to assess the model's effectiveness in classifying and detecting network anomalies.

2.2.1. Dataset

In this work, the models are trained using a Dataset called "IoT Modbus". This dataset and the source code are present in the github repository provided by [Bai 2023]. The IoT Modbus dataset was specifically designed for analyzing network traffic in IoT environments, focusing on industrial communication protocols. It contains a variety of labeled

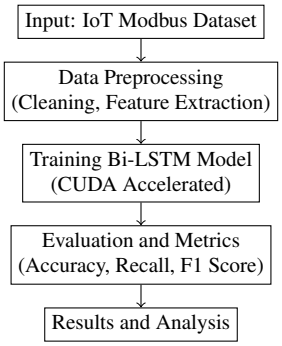

Figure 1. Workflow of the proposed Bi-LSTM approach.

samples, representing both normal and anomalous behaviors within Modbus communication flows. These anomalies simulate potential security threats, such as unauthorized access, command injection, and communication disruptions.

The dataset consists of 287,194 records, each containing 7 features that capture various aspects of network traffic, including data such as timestamps on IoT device telemetry and other attributes specific to Modbus communication, such as values returned by Modbus read functions. The dataset is structured to include both normal traffic and anomalous events, making it highly suitable for anomaly detection experiments. One of the feature columns is the *Label* field, which numerically represents whether a traffic instance is normal (0) or an anomaly (1), which is useful for cases where binary analysis is desired. The dataset includes a categorical field *type*, which specifies the type of attack present in the network traffic. This field represents a tag for attack categories (the classes), such as normal, DoS, DDoS, and backdoor attacks, and normal records.

2.2.2. Experiment Execution

The proposed approach is evaluated against a baseline "Lightweight LSTM" model (without CUDA optimization) provided by [Bai 2023]. Experiments were executed in a Kaggle environment equipped with an NVIDIA P100 GPU. Preprocessing involved timestamp unification, feature extraction, missing value handling, and data type casting. Model hyperparameters, including hidden layer sizes and learning rates, were selected through empirical grid-search optimization, tracking validation loss stabilization. The training process utilized the Adam optimizer with gradient clipping to prevent gradient explosion and specialized class-weighting coefficients to address dataset imbalance. Finally, model evaluation focused on assessing effectiveness in detecting positive anomalous samples across the test set using standard classification metrics.

3. Results

This section presents the performance evaluation of both the Lightweight-LSTM and the proposed CUDA-accelerated Bidirectional LSTM (Bi-LSTM) models. The analysis is grounded on standard classification metrics: Accuracy, Precision, Recall, and F1 Score, which collectively offer a view of model behavior, particularly in the context of anomaly detection in industrial IoT traffic.

Tables 1 and 2 summarize the results, showing that Bi-LSTM substantially outper-

forms the baseline across all metrics with 99.96% accuracy, 99.76% recall, and 99.68% F1 Score. Regarding efficiency, the proposed model required 9.69 seconds per epoch (≈15 min total) under CUDA acceleration on the NVIDIA P100 GPU, compared to 6.42 seconds per epoch for the baseline. Although bidirectional modeling doubles the recurrent workload (processing forward and backward states), parallel execution limited the execution overhead to a 1.51× factor, validating its high-throughput viability.

Table 1. Lightweight-LSTM Evaluation Metrics

Metric	Value
Accuracy	0.73
Precision	0.32
Recall	0.89
F1 Score	0.47

Table 2. Bidirectional-LSTM Evaluation Metrics

Metric	Value
Accuracy	0.9996
Precision	0.9996
Recall	0.9976
F1 Score	0.9968

Figures 2 and 3 illustrate the evolution of training loss and accuracy over epochs for both models. The Lightweight-LSTM (Figure 2) exhibits less stability during training, with fluctuations in both metrics that suggest a lower generalization capacity. In contrast, the Bi-LSTM model (Figure 3) demonstrates a faster convergence, with the loss decreasing consistently and the accuracy curve stabilizing at higher values in fewer epochs. This behavior reflects the architectural advantages of bidirectional learning and the performance benefits conferred by CUDA-based GPU acceleration in PyTorch.

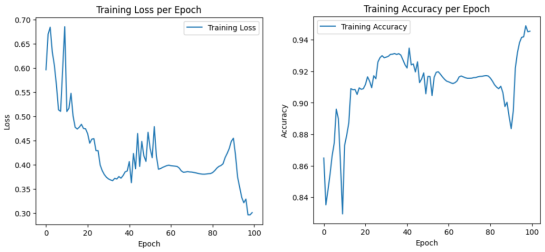

Figure 2. Lightweight LSTM Curves

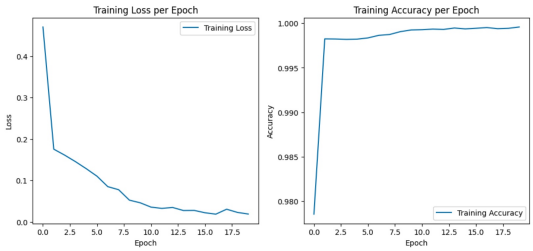

Figure 3. Bidirectional LSTM Curves

Figure 4 presents the confusion matrix for the Bi-LSTM model, offering a detailed insight into the per-class classification performance. The dominance of high values along the diagonal signifies excellent class discrimination, with negligible misclassification across categories such as *normal*, *backdoor*, *injection*, *scanning*, and others. This confirms the model's robustness and reliability when applied to heterogeneous network traffic in industrial IoT contexts.

In summary, the empirical findings validate the efficacy of the proposed architecture. The integration of bidirectional temporal modeling with GPU-accelerated training enhances detection performance and enables efficient learning in large-scale network datasets.

4. Conclusion and Future Works

This work proposed a Bidirectional LSTM-based approach for anomaly and attack detection in network traffic. The model effectively captured temporal dependencies in both directions, yielding strong classification performance. Its capabilities suggest broader applications, such as forecasting network behavior and optimizing IoT traffic. However, lim-

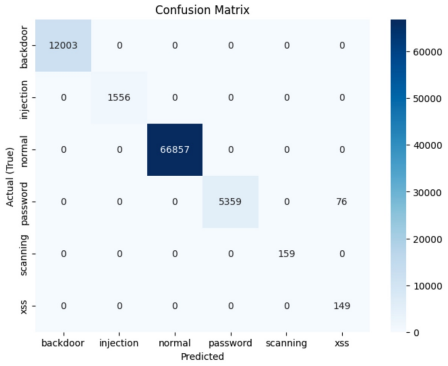

Figure 4. Confusion Matrix of the Bi-LSTM Model

itations include reliance on a single Modbus dataset and offline training, which may hinder generalization and real-time deployment. Future work will address diverse datasets, real-time inference, ensemble strategies, and adaptive learning to enhance robustness in dynamic environments.

References

Abbas, S., Alsubai, S., Ojo, S., Sampedro, G., Almadhor, A., Hejaili, A., and Bouazzi, I. (2024). An efficient deep recurrent neural network for detection of cyberattacks in realistic iot environment. *The Journal of Supercomputing*, 80:1–19.

Azab, A., Khasawneh, M., Alrabaee, S., Choo, K.-K. R., and Sarsour, M. (2024). Network traffic classification: Techniques, datasets, and challenges. *Digital Communications and Networks*, 10(3):676–692.

Bai, H. (2023). Iot-anomaly-detection. Accessed: 2024-05-17.

Gale, T., Zaharia, M., Young, C., and Elsen, E. (2020). Sparse gpu kernels for deep learning. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, SC '20. IEEE Press.

Keirsbilck, M. V., Keller, A., and Yang, X. (2019). Rethinking full connectivity in recurrent neural networks. *ArXiv*, abs/1905.12340.

Rajasinghe, N., Samarabandu, J., and Wang, X. (2018). Insecs-dcs: A highly customizable network intrusion dataset creation framework. In *2018 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*.

Rizi, M. H. P. and Seno, S. A. H. (2022). A systematic review of technologies and solutions to improve security and privacy protection of citizens in the smart city. *Internet of Things by Elsevier*, 20:100584.

Tayfour, O. E., Mubarakali, A., Tayfour, A. E., Marsono, M. N., Hassan, E., and Abdelrahman, A. M. (2023). Adapting deep learning-lstm method using optimized dataset in sdn controller for secure iot. *Soft Computing*.

Yang, M., Moon, J., Yang, S., Oh, H., Lee, S., Kim, Y., and Jeong, J. (2022). Design and implementation of an explainable bidirectional lstm model based on transition system approach for cooperative ai-workers. *Applied Sciences (Switzerland)*, 12(13).

Zhao, Y. (2023). Complete guide to rnn, lstm, and bidirectional lstm. last accessed: 2024-05-23.