An Intrusion Detection System using a Hybrid Lightweight Deep Learning Algorithm

Rusul H. Altaie
College of Information Technology, Department of Software, University of Babylon, Iraq
rusul.jasem@uobabylon.edu.iq (corresponding author)

Haider K. Hoomod
College of Education, Department of Computers, Mustansiriyah University, Iraq
drjnew@gmail.com

Received: 27 April 2024 | Revised: 12 June 2024 | Accepted: 15 June 2024

ABSTRACT

Cybercriminals are interested in the Internet of Things (IoT) more than ever due to its remarkable growth pace. This assertion is supported by the growing number of cyberattacks on IoT devices and intermediary communication mediums. IoT attacks that go unnoticed for a long time can result in serious service disruptions and monetary losses. Real-time intrusion detection on IoT devices is necessary to ensure the dependability, security, and profitability of IoT-enabled services. This study proposes a lightweight deep-learning method for detecting intrusions in IoT devices. The proposed system uses a hybrid Convolution Neural Network (CNN) with Long Short Term Memory (LSTM). Three distinct models, CNN, LSTM, and the proposed hybrid CNN+LSTM were used to identify intrusions in the UNSW-NB15 dataset. The proposed hybrid model was used to identify event characteristics on a Raspberry Pi3 device. To minimize computation costs, CNN and LSTM were stacked without the final layer to maximize convergence speed. CNN and LSTM layers are mapped to the sample marker space using fully linked layers and a softmax layer. The average accuracy, recall, precision, and F1-score of the proposed hybrid intrusion detection system were 98.78%, 98.09%, 97.88%, and 97.99%, respectively.

Keywords-Convolutional Neural Network (CNN); Deep Learning (DL); Intrusion Detection System (IDS); Internet of Things (IoT); Long Short-Term Memory (LSTM)

I. INTRODUCTION

IoT devices are used in many and various fields, such as manufacturing, transportation, smart homes, healthcare, and blockchain monitoring. Network system administrators use Network Intrusion Detection Systems (NIDS) to detect different types of security breaches within an organization's network [1-3]. An IDS is a kind of security system that can monitor network traffic and spot unusual or hostile activities. Technically, an IDS is just the classification problem of determining whether any given network activity is normal or abnormal. Binary classification and multiclass classification are two types of classification. When classifying binary data, the system's output is attack or normal. However, attack types can also be identified by a multiclass categorization. This study uses multiclassification to detect intrusions.

Deep learning has meteoric growth in the arena of machine learning. This study proposes a hybrid deep learning method to create an IDS for IoT utilizing Long Short-Term Memory (LSTM) and Convolution Neural Network (CNN) approaches to increase efficiency and effectiveness. Deep learning approaches have been deployed to increase detection accuracy and overcome some of the limitations of traditional detection methods [3-5]. However, processing large amounts of data requires specialized expertise. When network traffic enters or leaves an organization's network equipment, a NIDS records it, examines it, and alerts in case of detecting an attack. NIDS are divided into two varieties, depending on intrusion detection methods: i) Signature-based NIDS (SNID) and ii) Anomaly Detection-based NIDS (ADNIDs). Anytime a variation from the typical traffic pattern is noticed, an ADNIDS labels the network traffic as an intrusion. In contrast, SNIDSs have low false alarm rates, excellent detection accuracy, and are useful in identifying known threats.

This study developed and compared a hybrid CNN-LSTM and single CNN and LSTM approaches. In comparison to LSTM or CNN findings, the proposed hybrid model achieved greater accuracy rates for intrusion detection. The development of LTSM addressed the vanishing gradient issue that conventional Recurrent Neural Networks (RNNs) faced. In contrast to other RNNs, such as hidden Markov models and other sequence learning techniques, LSTMs are less sensitive to gap length. The goal is to enable LSTM or short-term memory that can transcend thousands of time steps [6-10]. CNNs are some of the most widespread and extensively used neural networks. CNNs have a role in deep learning's recent surge in popularity. The main advantage of CNNs is that they can automatically detect essential things without human supervision. For self-learning, CNNs employ filter (or kernel) optimization. Regularized weights are applied over fewer connections, avoiding the backpropagation problems that earlier neural networks had with growing and vanishing gradients. A CNN consists of three types of layers: an Output Layer (OL), Hidden Layers (HL), and an Input Layer (IL). The hidden layers of a CNN contain one or more convolution-performing layers. This typically comprises a layer that creates the convolution kernel's dot product using the layer's input matrix and ReLU as an activation function [2, 11].

In [13], a dataset was presented for complicated Industrial IoT (IIoT) networks, called X-IIoTID, which includes data from connected devices from various manufacturers and platforms. Many different methods have been proposed to protect against ransomware attacks on Industrial IoT (IIoT) network devices. In [14], an autoencoder was used in the initial cleanup of the two-section model data to provide the best representation. This facilitated the identification and detection of intrusions in the deep neural network. Then, several tests were performed using the proposed model with datasets such as NSL-KDD, ISOT, and X-IIoTID. According to the study findings, ransomware intended for IoT devices could have been detected with a high probability. The security of an IoT network can be ensured and improved in several ways. In [15], an IDS was proposed using deep learning. This study examined LSTM, Random Neural Networks (RandNN), and Feed-Forward Neural Networks (FFNN). Intrusion detection accuracy was similar between models: 96.93% for FFNN, 96.85% for LSTM, and 96.42% for RandNN. In [16], a general and comprehensive method was proposed for IoT intrusion detection, employing a bidirectional LSTM, which exceeded 95% accuracy in detecting attacks.

II. METHODOLOGY

NIDSs describe activities as hostile and identify violations related to public policies. In addition, they look for security flaws or malicious activity on a system or network. IDSs work by looking for indicators of known assaults or behavioral abnormalities. This study employed lightweight deep learning techniques in two phases: event detection, based on a hybrid CNN and LSTM (HCNN-LSTM) without an output layer, and intrusion detection using the output layer of the HCNN-LSTM, as shown in Figure 1. The UNSW-NB15 dataset identifies normal and aberrant data, containing nine attack types: DoS, worms, backdoors, fuzzers, analysis, exploits, generic, reconnaissance, and shellcode. From this dataset, 175,341 records were used for training and 82,332 records were used for testing. The proposed approach works in two phases: event detection and intrusion detection. The hybrid CNN+LSTM was used to detect events (attack, normal) in the data, using the algorithms in an overlapping manner. At first, the CNN algorithm is used, which gives a set of features that are fed into the LSTM. Then, these features are encrypted by a hybrid encryption algorithm (PRESENT+SPECK) and CNN is used with the output layer using the softmax function to detect intrusion and perform classification.

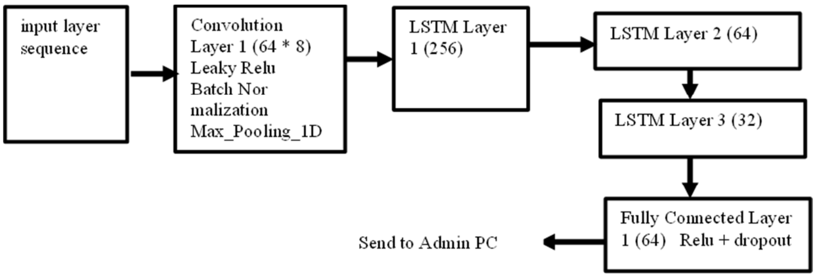
input layer sequence
Convolution Layer 1 (64 * 8) Leaky Relu Batch Nor malization Max_Pooling_1D
LSTM Layer 1 (256)
LSTM Layer 2 (64)
LSTM Layer 3 (32)
Fully Connected Layer 1 (64) Relu + dropout
Send to Admin PC

Fig. 1. The proposed network design for anomaly detection.

In the CNN method, 1D_convolution for a linear activation layer, and a maximum pooling layer make up each convolution layer. The network is also subjected to a batch normalization layer to exploit the convergence rate. To ensure that the model can learn enough features, LSTM layers are employed. The CNN and LSTM layers provide highly abstract features, which are then mapped to the sample marker space using a fully linked layer. Table I shows the parameters used in the experiments for event detection.

To facilitate comparisons, a 1-D CNN and a pure LSTM were trained with identical parameters. To identify event characteristics in a Raspberry Pi 3, the hybrid lightweight CNN+LSTM without an output layer was used to minimize the computational requirements. The HCNN+LSTM was layered with the last layer to detect intruders in the network, as shown in Figure 2. The same hyperparameters were used for the intrusion detection. Figure 3 shows the complete intrusion detection architecture.

TABLE I. HYPERPARAMETERS USED FOR EVENT DETECTION

Parameter	Value
Number of epochs	30
Batch size	32
Learning rate	0.001
Dropout	0.3
No. of neurons in the dense layer	1024

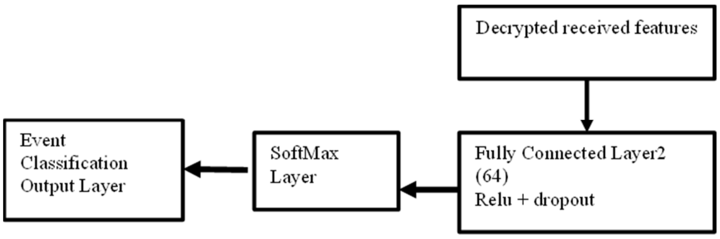
Decrypted received features
Fully Connected Layer2 (64) Relu + dropout
SoftMax Layer
Event Classification Output Layer

Fig. 2. Block diagram of intrusion detection.

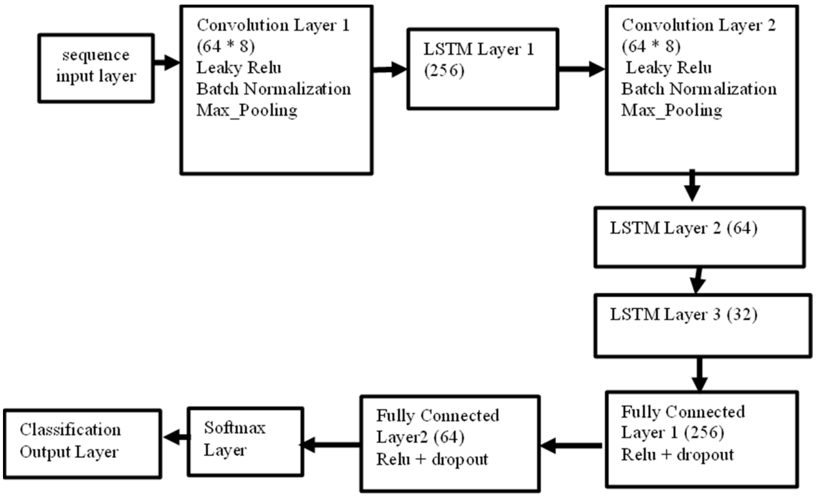
sequence input layer
Convolution Layer 1 (64 * 8) Leaky Relu Batch Normalization Max_Pooling
LSTM Layer 1 (256)
Convolution Layer 2 (64 * 8) Leaky Relu Batch Normalization Max_Pooling
LSTM Layer 2 (64)
LSTM Layer 3 (32)
Fully Connected Layer 1 (256) Relu + dropout
Fully Connected Layer2 (64) Relu + dropout
Softmax Layer
Classification Output Layer

Fig. 3. Proposed architecture for intrusion detection

III. RESULTS AND DISCUSSION

The proposed approach uses 1D_convolutional layers for feature extraction. This convolutional layer avoids the gradient disappearance issue by utilizing the Leaky ReLu activation function, as it performs better for 1D time series. A maximum pooling layer of length two comes after each convolution layer to shrink the feature vector and provide higher-level abstract feature information to the network that follows. Better results are obtained using LSTM by including a memory gate into the RNN to address the sequence's long-term dependence issues. As shown in Table II, the experimental results show that the CNN-LSTM network had more accuracy than the 1D CNN and LSTM. The experimental results of the event detection system compared to CNN and LSTM for multiclassification are shown in Table III.

TABLE II. EXPERIMENTAL RESULTS OF THE EVENT DETECTION SYSTEM COMPARED WITH CNN AND LSTM

Model	Accuracy (%)	Recall (%)	Precision (%)	F1 score (%)
CNN-LSTM	96.78	97.44	96.09	96.49
CNN	94.12	95.76	94.96	95.41
LSTM	95.10	95.42	94.70	96.01

The system recognizes and classifies many types of attacks, including teardrop, buffer overflow, rootkit, and others, based on the headers. There is also an unnamed category, and the packet is marked unknown if the IDS cannot classify it under any known attack category. On the other hand, the IDS is based on the features of a normal packet. It evaluates a newly received packet and compares its attributes to those of a standard one. The packet is labeled as an anomaly if the difference is greater than a particular threshold. Otherwise, it is classified as a typical packet. Table III shows the experimental results of the proposed NIDS compared with CNN and LSTM for multiclassification. The findings in Tables II and III show that the proposed approach outperforms the simple algorithms in all performance metrics.

TABLE III. EXPERIMENTAL RESULTS OF PROPOSED NIDS

Model	Accuracy (%)	Recall (%)	Precision (%)	F1 score (%)
HCNN-LSTM	98.78	98.09	97.88	97.99
CNN	95.90	96.78	97.31	96.60
LSTM	95.12	95.65	96.87	96.78

IV. CONCLUSION

This study introduced a hybrid LSTM+CNN model for intrusion detection in IoT environments. In the training phase, the proposed approach can efficiently extract certain features from the dataset. Two deep learning algorithms, CNN and LSTM, were used to detect events (attack, normal) in the data in an overlapping manner. First, a CNN algorithm is used to provide a set of features that are fed into the 3-layer LSTM model. The LSTM outputs become inputs to a CNN without applying an output layer. These features are then encrypted by hybrid encryption algorithms (PRESENT+SPECK), and then a CNN is used, with the output layer using the softmax function to detect intrusion and perform classification. This approach is crucial for identifying features in network traffic that are involved in anomalous incursions and differentiating between anomalous and regular traffic. The UNSW-NB15 network intrusion dataset was used to evaluate the accuracy of the proposed model. In testing, the proposed NIDS outperformed the single LSTM and CNN approaches in terms of normal/anomaly detection. The proposed model demonstrated a high degree of accuracy in identifying attack traffic. Future studies will examine the performance of the proposed method on different types of data to identify intrusions.

REFERENCES

[1] D. K. Singh and M. Shrivastava, "Evolutionary Algorithm-based Feature Selection for an Intrusion Detection System," *Engineering, Technology & Applied Science Research*, vol. 11, no. 3, pp. 7130–7134, Jun. 2021, https://doi.org/10.48084/etasr.4149.

[2] M. Anwer, S. M. Khan, M. U. Farooq, and Waseemullah, "Attack Detection in IoT using Machine Learning," *Engineering, Technology & Applied Science Research*, vol. 11, no. 3, pp. 7273–7278, Jun. 2021, https://doi.org/10.48084/etasr.4202.

[3] R. H. Altaie and H. K. Hoomod, "Artificial Intelligent Management for Internet of Things: A Review," in *2022 4th International Conference on Current Research in Engineering and Science Applications (ICCRESA)*, Baghdad, Iraq, Dec. 2022, pp. 179–184, https://doi.org/10.1109/ICCRESA57091.2022.10352510.

[4] Y. Imrana, Y. Xiang, L. Ali, and Z. Abdul-Rauf, "A bidirectional LSTM deep learning approach for intrusion detection," *Expert Systems with Applications*, vol. 185, Dec. 2021, Art. no. 115524, https://doi.org/10.1016/j.eswa.2021.115524.

[5] A. Awajan, "A Novel Deep Learning-Based Intrusion Detection System for IoT Networks," *Computers*, vol. 12, no. 2, Feb. 2023, Art. no. 34, https://doi.org/10.3390/computers12020034.

[6] S. Choudhary and N. Kesswani, "Analysis of KDD-Cup'99, NSL-KDD and UNSW-NB15 Datasets using Deep Learning in IoT," *Procedia Computer Science*, vol. 167, pp. 1561–1573, Jan. 2020, https://doi.org/10.1016/j.procs.2020.03.367.

[7] P. More and P. Mishra, "Enhanced-PCA based Dimensionality Reduction and Feature Selection for Real-Time Network Threat Detection," *Engineering, Technology & Applied Science Research*, vol. 10, no. 5, pp. 6270–6275, Oct. 2020, https://doi.org/10.48084/etasr.3801.

[8] R. Alsulami, B. Alqarni, R. Alshomrani, F. Mashat, and T. Gazdar, "IoT Protocol-Enabled IDS based on Machine Learning," *Engineering, Technology & Applied Science Research*, vol. 13, no. 6, pp. 12373–12380, Dec. 2023, https://doi.org/10.48084/etasr.6421.

[9] S. Smys, A. Basar, and H. Wang, "Hybrid Intrusion Detection System for Internet of Things (IoT)," *Journal of IoT in Social, Mobile, Analytics, and Cloud*, vol. 2, no. 4, pp. 190–199, Sep. 2020, https://doi.org/10.36548/jismac.2020.4.002.

[10] B. K. Park and C. J. Kim, "Unsteady Heat Flux Measurement and Predictions Using Long Short-Term Memory Networks," *Buildings*, vol. 13, no. 3, Mar. 2023, Art. no. 707, https://doi.org/10.3390/buildings13030707.

[11] M. A. Alsoufi *et al.*, "Anomaly-Based Intrusion Detection Systems in IoT Using Deep Learning: A Systematic Literature Review," *Applied Sciences*, vol. 11, no. 18, Jan. 2021, Art. no. 8383, https://doi.org/10.3390/app11188383.

[12] Y. Slimani and R. Hedjam, "A Hybrid Metaheuristic and Deep Learning Approach for Change Detection in Remote Sensing Data," *Engineering, Technology & Applied Science Research*, vol. 12, no. 5, pp. 9351–9356, Oct. 2022, https://doi.org/10.48084/etasr.5246.

[13] M. Al-Hawawreh, E. Sitnikova, and N. Aboutorab, "X-IIoTID: A Connectivity-Agnostic and Device-Agnostic Intrusion Data Set for Industrial Internet of Things," *IEEE Internet of Things Journal*, vol. 9, no. 5, pp. 3962–3977, Mar. 2022, https://doi.org/10.1109/JIOT.2021.3102056.

[14] M. Al-Hawawreh, E. Sitnikova, and N. Aboutorab, "Asynchronous Peer-to-Peer Federated Capability-Based Targeted Ransomware Detection Model for Industrial IoT," *IEEE Access*, vol. 9, pp. 148738–148755, 2021, https://doi.org/10.1109/ACCESS.2021.3124634.

[15] Z. Xu, Y. Guo, C. Chakraborty, Q. Hua, S. Chen, and K. Yu, "A Simple Federated Learning-Based Scheme for Security Enhancement Over Internet of Medical Things," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 2, pp. 652–663, Oct. 2023, https://doi.org/10.1109/JBHI.2022.3187471.

[16] M. Shafiq, Z. Tian, A. K. Bashir, X. Du, and M. Guizani, "IoT malicious traffic identification using wrapper-based feature selection mechanisms," *Computers & Security*, vol. 94, Jul. 2020, Art. no. 101863, https://doi.org/10.1016/j.cose.2020.101863.

AUTHORS PROFILE

Rusul Haider Altaie is an Associate Lecturer in the College of Information Technology, University of Babylon, Iraq. She holds an M.Sc. degree in Computer Science with a specialization in image analysis. Her research areas are image/signal processing and medical image analysis. Her research interests include image/signal processing, biometrics, medical image and analysis, and the Internet of Things.

Haider K. Hoomod has a B.Sc. in Electrical Engineering from the Electrical Engineering Department, at Technology University in Baghdad, Iraq. He has an M.Sc. and Ph.D. in Computer Networking from the Communication Engineering Department, Technology University, Baghdad. He specializes in network engineering, wireless communications, IoT and WoT management and security, image processing, network artificial intelligence, artificial neural networks, and network security. He is a publisher and reviewer of many engineering papers across the world in IEEE and many scientific committees and journals and member of many scientific committees. He is also the supervisor of many M.Sc. and Ph.D. students in many Iraqi universities.