DeepCyber-IDS: a Deep Learning Based Intrusion Detection System

Zaman Najii Hussein
Computer Techniques Engineering Department, Electrical Engineering Technical College
Middle Technical University
Baghdad, Iraq
bbc0079@mtu.edu.iq

Dalal Abdulmohsin Hammood
Cybersecurity Technology Engineering Department, Electrical Engineering Technical College
Middle Technical University
Baghdad, Iraq
dalal.hammood@mtu.edu.iq

Ziad Qais Al-Abbasi
Electromechanical Techniques Department, Baqubah Technical College (BTC),
Middle Technical University
Diyala, Iraq
ziad.al-abbasi@mtu.edu.iq

***Abstract*— Due to the rise in advanced and more frequent cyberattacks, traditional IDS systems are unable to identify new threats and can create a lot of false alarms. In this paper, a new IDS called DeepCyber-IDS is discussed, which addresses these issues by using deep learning to spot complex network traffic patterns and require less fine-tuning or creation of new features. The system combines CNN, GRU, and LSTM layers to draw useful information from the raw traffic data in terms of space and time. Unlike legacy IDS systems, DeepCyber-IDS learns on its own from unrefined input, meaning it can fit into various network situations without needing to be guided by outside features. It was tested on three commonly used datasets to see how well it performed. NSL-KDD, UNSW-NB15, and SmartGrid. The model got a perfect F1-Score and a 100% precision and recall rate on UNSW-NB15 and SmartGrid, and an accuracy score of 99.84% on NSL-KDD. The processing times varied from 469 milliseconds to about 53 seconds, meaning it can be used in real-time monitoring situations.They reveal that DeepCyber-IDS can protect against a wide array of threats and is also capable of being easily improved over time.**

Keywords— cyberattacks; intrusion detection system; deep learning; LSTM; CNN

I. INTRODUCTION

With more digital technology appearing everywhere, there are more opportunities for cyberattacks. Since modern networks are made up of different components and grow all the time, they become more susceptible to sophisticated attacks that go unnoticed [1]. Traditional IDS solutions both with signature-based detection and based on watching for anomalies have not been able to respond well to advanced persistent threats, polymorphic malware, and vulnerabilities that first appear on the same day. Because they often have high false positives, they can make security teams' work much harder and slow down operations. The Internet of Things (IoT) involves a wide variety of interconnected gadgets that can send data and services between themselves, without humans getting involved [2], . Even though automation and efficiency got better over time, there has also been an increase in Distributed Denial of Service (DDoS) attacks. Along with saturating computer resources and internet bandwidth, some DDoS attacks these days also use extortion, adding to their complexity in being handled [3].

When a lot of data moves across a network rapidly, real-time detection of intrusions gets more problematic. Always remember to find the right balance between detecting threats and working efficiently and quickly. Isolating unusual activity on networks is difficult because there is so much data and it is so variable and noisy. The addition of wormhole, sinkhole, flooding, and jamming complicates things by bringing new ways to disrupt communication while also making it look as if real users are communicating normally [5], [4]. Having extra or unnecessary information in traffic data delays the classifier in telling normal activity apart from malicious activity, demands more time and effort, and leads to a high rate of false indications [5]. To get around these issues, there has been a trend to handle intrusion detection as a classification challenge, with ML and DL techniques used more widely as a result [6]. K-Nearest Neighbors (KNN), Decision Trees (DT), Support Vector Machines (SVM), and Random Forests (RF) are often used as they are both understandable and require only moderate amount of resources. In addition, recent research has revealed that DL networks called autoencoders, deep belief networks, and convolutional neural networks (CNNs) work better at finding complex patterns and modeling large data sets [7],[8].

The method suggested here advances the field of intelligent and scalable IDS, ready to be used in today's IoT and cyber-physical systems.To fix these issues, the research presents DeepCyber-IDS, a new IDS framework based on deep learning. It is important to create a design that can spot changes well and scale and perform well on various types of networks. DeepCyber-IDS focuses on being both accurate and suited to various environments, catching both already known and new types of attacks with very little user or developer involvement.

A combination of a CNN-GRU-LSTM model and a special network scheme is used to strengthen identifying DDoS and different network intrusions. The use of optimization enhances the model's accuracy in detection, its ability to reduce features, and the speed at which the model executes. By comparing the model's results with those from DeepCyber-IDS, it is clear that the enhanced model is faster and more applicable in real-time.

II. LITERATURE REVIEW

With the use of deep learning and feature selection, IDS for IoT and cyber-physical systems have made significant improvements. Because of Mohy-eddine et al. [9], designed IDSs for IoT networks now use an integrated feature selection method and KNN algorithm . They looked at different ways to do feature selection, including using things like Genetic Algorithms, simple statistics, and Principal Component Analysis. They found that choosing the top ten important features greatly improves the detection process. An other hybrid IDS model was introduced by a team of researchers, using a mix of both deep and shallow learning algorithms [10]. They used the SMOFS method to select related parameters among a large number of them. A

Siamese Neural Network was applied to improve how well the features distinguish between data points, and it also handled changes in the shape of the intrusion data more effectively.

According to Syed et al. [11], the next-generation architecture they outlined for an IDS in fog-cloud IoT works by filtering data based on profiles of attackers in a distributed manner. Still, it did not have a unified tool for choosing features, so it could not reach high optimization levels with large applications. A comprehensive review by Chen et al. [12] looked at different ways to protect the Internet of Medical Things (IoMT) and also checked out several types of artificial intelligence systems used for this purpose. According to the study, there is a greater need for advanced AI to handle cybersecurity and ethics problems, especially in cloud-fog-edge systems.

Binbusayyis et al. [13] also came up with a way to protect healthcare IoMT systems from data pollution by using a method that works together with the main system but keeps sensitive data away from poisoning. The combination of TFL and clustering in their model made it robust to various manipulation attacks without losing the integrity of their model. In looking at traditional ML models for IoMT security, one study [14] tested KNN, Naïve Bayes, SVM, ANN, and Decision Trees using the Bot-IoT dataset to see how well they work. It was found that using classical classifiers, and more specifically XGBoost, provided 100% accuracy when detecting malicious traffic.

Jithish et al. [15] published a novel IDS designed for Cyber-Physical Manufacturing Systems (CPMS), which is based on Kernel PCA (KPCA) and Self-Organizing Maps (SOMs). Their testing using the CSTR simulation model showed they caught almost every problem, doing better than most older ways of checking for security issues. Maseno and Wang designed an improved IDS that combines ELM with SVM [16]. They selected the features by using a genetic algorithm and then continued with sequential forward selection, running their model on the IoT-ToN and UNSW-NB15 datasets. As a result, the model performed better, with 99% and 86% accuracy, supporting the usefulness of using hybrid feature selection in IDS. IDS performance was improved by Fadhil et al. [17] through the implementation of the Lion Optimization Algorithm (LOA) and Grey Wolf Optimizer (GWO) on IDS systems. As part of their study, they developed a new IDS that combines LOFS and GWM to lessen the number of features in a CNN-LSTM deep learning framework. Using benchmark datasets such as NSL-KDD, their system managed to achieve an accuracy of over 99.26%, which led to more exact detection of anomalies and fewer false positives.

III. Prepare Methodology

DeepCyber-IDS is built using a mix of CNN, GRU, and LSTM so that it can efficiently understand and classify network traffic as illustrated in Figure 1. It has multiple layers that help capture how different parts of the network traffic change over both space and time. CNN layers help find features in small sections of data, GRUs make sure the model understands the short patterns that happen near each other, and LSTM units help the model pick up on longer patterns across the sequence. This way, DeepCyber-IDS is capable of finding many different attack patterns, including simple and quick anomalies as well as elaborate intrusions.

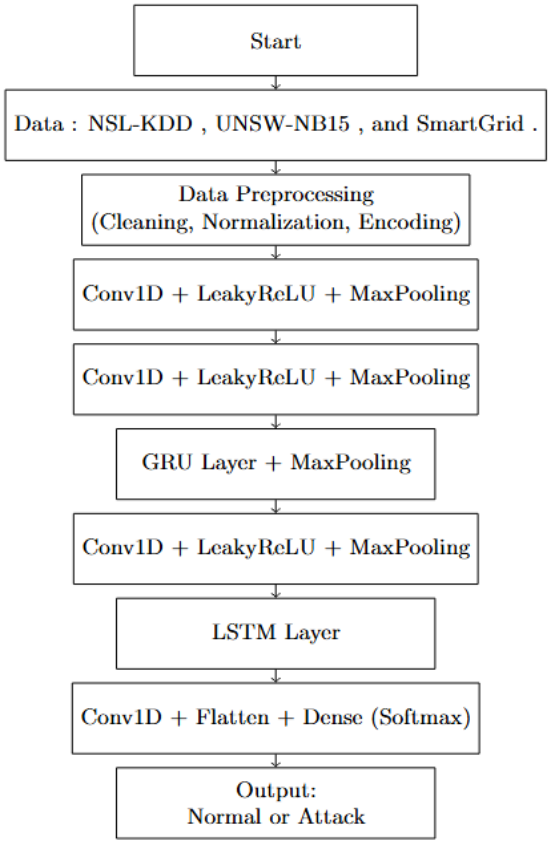

Fig. 1. Workflow of DeepCyber-IDS Model Architecture

A. *Dataset Description*

The model was trained and tested with three sets of data that most experts think are reliable: NSL-KDD [18], UNSW-NB15 [19], and SmartGrid [20]. An advantage of the NSL-KDD over the KDD'99 dataset is that less number of samples are needed due to its balanced and refined state. The UNSW-NB15 dataset contains behaviors that modern networks might see and more types of attack patterns, helping make it useful for testing how well machines and networks handle today's security challenges. On the other hand, SmartGrid conducts experiments by simulating attacks on industrial control systems and smart grids based on the Industrial Internet of Things (IIoT). All datasets were cleaned up by getting rid of noisy data, making sure all the numbers in the data were in a similar range, and turning categorical data into numbers that the model could understand.

B. *Model Architecture*

The main structure of DeepCyber-IDS is built to understand patterns and trends in data received one after another from the network as shown in Table (1). The model starts with a 1D Convolutional Neural Network (Conv1D) layer where it uses 8 filters of size 3 and moves along the input sequence one step at a time. This part of the process allows for the detection of sequence patterns within ports and protocol headers. LeakyReLU is used to activate the output with alpha at 0.3, so that very negative numbers do not cause the neurons to become inactive. The next layer, MaxPooling1D, uses a pool size of 2 and same padding, which helps downsize the feature map without reducing important patterns.

Then, more convolutional layers are added, with each one taking the depth to 16 and then 32 filters, and accompanied by LeakyReLU and pooling. It allows the model to take raw

information and extract more significant features about how attacks occur. The extracted features are then sent into a Gated Recurrent Unit (GRU) layer consisting of 16 units. GRUs help in catching short-term patterns as well as activities that happen all at once or occur in a way out of the norm for the sequence as a whole. A second MaxPooling1D operation is also used to squeeze the time-based information even more.

This is followed by a LSTM layer with 16 units, which does a good job at picking up patterns that happen far apart, like attack attempts that come long after something started, or cases where someone stays online for a long time, which ducks might be a sign of someone trying to break in. The LSTM is important because it can pick up patterns and connections in the sequence of words or numbers that regular models might miss. After finishing the convolutional layers, there is another layer with 16 filters, which is followed by squashing the sequence into 2 output variables through a Dense layer and using softmax to perform the final normal vs. attack classification. Thanks to the combination of CNN-GRU-LSTM, DeepCyber-IDS is capable of picking up complex features across time and space, making it useful for identifying ongoing and recent cyber threats without delay.

TABLE I. DEEPCYBER-IDS MODEL ARCHITECTURE

Layer Type	Output Shape	Parameters
Conv1D	(None, 77, 8)	32
LeakyReLU	(None, 77, 8)	0
MaxPooling1D	(None, 39, 8)	0
LeakyReLU	(None, 39, 8)	0
Conv1D	(None, 39, 16)	400
MaxPooling1D	(None, 39, 16)	0
Conv1D	(None, 39, 16)	784
LeakyReLU	(None, 39, 16)	0
MaxPooling1D	(None, 39, 16)	0
Conv1D	(None, 39, 32)	1568
LeakyReLU	(None, 39, 32)	0
MaxPooling1D	(None, 39, 32)	0
GRU	(None, 39, 16)	2352
LeakyReLU	(None, 39, 16)	0
MaxPooling1D	(None, 39, 16)	0
Conv1D	(None, 39, 64)	3136
LeakyReLU	(None, 39, 64)	0
MaxPooling1D	(None, 20, 64)	0
Conv1D	(None, 20, 32)	6176
LeakyReLU	(None, 20, 32)	0
LSTM	(None, 20, 16)	3136
Conv1D	(None, 20, 16)	784
Flatten	(None, 320)	0
Dense	(None, 2)	642

C. *Training and Evaluation*

The way DeepCyber-IDS was trained aimed to achieve good detection results and work effectively on different data sets. The optimizer used in the model, Adam, allows for an adaptive learning rate, making it efficient in training networks with constantly altering gradients. The loss function used was categorical cross-entropy, which works well for problems with many classes, and also works just as well when classifying binary data if you use one-hot coding for the labels.

Training applied mini-batch gradient descent, with both the batch sizes and learning rates picked using a grid search. The study tested batch sizes of 16, 32, and 64, and learning rates between 104 -10 −4 and 102 -10 −2. While the maximum number of epochs allowed was 100, the model stopped training early if the validation loss did not improve for ten consecutive epochs, so there was less risk of overfitting. Model validation was done using a separate part (usually 20%) of the training data to make sure that the performance results weren't affected by the way the model was built. Testing was carried out on the reserved part of the data, and five important metrics were checked. It checks the accuracy, precision, recall, F1-score, and also the inference time. Each dataset (NSL-KDD, UNSW-NB15, SmartGrid) was trained and checked on its own so that we could really compare how well the models work with different types of traffic and attacks. Because of this form of training, DeepCyber-IDS works well and can be relied on in real-world situations for detecting threats.

IV. RESULTS

To evaluate the performance of DeepCyber-IDS, experiments were done with three different sets of data, called benchmark datasets. NSL-KDD, UNSW-NB15, and SmartGrid as shown in Table 2. The model was independently assessed for each dataset by monitoring its outcome on five different metrics of performance. Precision, Recall, F1-score, Accuracy, and how fast the process takes.

TABLE II. MODEL PERFORMANCE METRICS

Dataset	Precision	Recall	F1-Measure	Accuracy	Time (Sec)
unsw_nb	1	1	1	1	53
nslkdd	1	1	1	0.9984	27
SmartGrid	1	1	1	1	469ms

On the UNSW-NB15 dataset, DeepCyber-IDS got 100% scores on all the evaluation metrics I mentioned. 1.00, Recall: 1.00, F1-score: 1.00, and Accuracy: 100% took around 53 seconds to finish. It proves that the model is good at handling the differences between different kinds of traffic, both legal and malicious, in modern networks. Results from using the NSL-KDD dataset were also outstanding, making Precision, Recall, and the F1-score 1.00 and the Accuracy 99.84%. It only took 27 seconds for the model to complete inference, which shows its ability to work with organized and large network data.

For the SmartGrid dataset, which deals with industrial IoT and smart grid situations, DeepCyber-IDS got exactly the right results when classifying attacks and it finished the test in about 0.469 seconds. The performance suggests that the model can adapt to different and quick-changing data datasets. All in all, these findings reflect that the DeepCyber-IDS architecture is reliable, runs well, and can be used in different network environments without needing prior feature or metaheuristic configuration.

V. DISCUSSION

The results obtained from checking DeepCyber-IDS on three different benchmark datasets show that the system works well on a variety of data sets, is accurate in detecting attacks, and can learn how to recognize new threats. The model did a really good job at correctly classifying email, always getting Precision, Recall, and F1-scores at 1.00, and Accuracy of 99.84% to 100% on the three email datasets it tested on. These metrics show that the system works well at telling normal traffic apart from bad traffic, which is really important for putting an IDS into use in real situations.

Compared to what others have said before, these results show that the analysis was done well and gave useful information. By way of example, models with classical machine learning algorithms (meaning Decision Trees, KNN, SVM) often demonstrate that higher accuracy comes at the cost of more false positives. Furthermore, methods that use different ways to select features usually need a lot of adjustments to work well, and they don't always do well when used with different datasets. In contrast, DeepCyber-IDS got its good results without needing feature selection or any optimization techniques, showing how well its in-built feature extraction works using CNN-GRU-LSTM layers.

One especially interesting result was how well the system did on the SmartGrid dataset, which has tough and complicated data from both electrical grids and industry. Achieving high accuracy in a short amount of time (469 ms) means the model can help detect intrusions in real-time on the internet of things and industry internet of things devices. This is important because these systems have to work quickly and correctly, or else bad things can start happening one after the other, or the system might just stop working.

However, there are a few weaknesses to it. The training data was made up of clean, standard datasets, meaning it has not yet been tested in realistic messy or intentional noisy environments. Additionally, even though the model works well, it still needs a lot of computing power when training, which can make it difficult to use on small devices that don't have much resources. The absence of a feature optimization step on purpose in this phase can also result in some input features that are unnecessary or don't add much value, so it might be a good idea to get rid of them to make the model work better.

VI. Conclusion

In this paper, the hybrid DeepCyber-IDS system was developed to help in detecting and classifying a variety of attacks on different kinds of networks. The use of CNN, GRU, and LSTM in the system allowed it to detect threats in network traffic using information about spatial layout and timing. Tests conducted on NSL-KDD, UNSW-NB15, and SmartGrid show that DeepCyber-IDS has small executive times, and close-to-100% accuracy. This reveals the strong and useful features of the architecture, both in different types of cybersecurity situations and in businesses.

This means DeepCyber-IDS was designed without relying on extra optimization tools, indicating it can directly process the input data and pick out significant patterns. Though the results are encouraging, future research will include the Lion Optimization Algorithm and other metaheuristic algorithms to improve both the efficiency and simplicity of the model. All in all, DeepCyber-IDS shows great promise for tackling current and future intrusion detection needs in both regular and industrial IoT systems.

References

[1] S. I. Ibrahim, D. A. Hammood, and L. H. Abed, "A comprehensive review of cancelable biometrics for cybersecurity solutions," The Fifth Scientific Conference For Electrical Engineering Techniques Research (EETR2024), vol. 3232, p. 020042, 2024, doi: 10.1063/5.0236371.

[2] S. I. Ibrahim, D. A. Hammood, and L. H. Abed, "Enhancing cloud data security with biometrics-based encryption and machine learning," International Journal of Advanced Technology and Engineering Exploration, vol. 12, no. 122, Jan. 2025, doi: 10.19101/ijatee.2024.111100624.

[3] S. T. A. Al-Latief, S. Yussof, A. Ahmad, S. M. Khadim, and A. Alkhayyat, "WAR Strategy Algorithm- based Hybrid Optimization for Accurate and Rapid Speech Recognition," Iraqi Journal for Computer Science and Mathematics, vol. 6, no. 1, Mar. 2025, doi: 10.52866/2788-7421.1243.

[4] D. A. Hammood, "A hybrid system based on machine learning and PSO for network intrusion detection," The Fifth Scientific Conference For Electrical Engineering Techniques Research (EETR2024), vol. 3232, p. 020041, 2024, doi: 10.1063/5.0236440.

[5] S. Kadhim, J. K. Siaw Paw, Yaw. C. Tak, S. Ameen, and A. Alkhayyat, "An Optimized Machine Learning Models by Metaheuristic Corona Virus Optimization Algorithm for Precise Iris Recognition," Advances in Artificial Intelligence and Machine Learning, vol. 05, no. 01, pp. 3389–3408, 2025, doi: 10.54364/aaiml.2025.51194.

[6] Saja Theab Ahmed, D. A. Hammood, R. F. Chisab, and Nurulisma Binti Hj. Ismail, "Medical Image Encryption and Decryption Based on DNA: A Survey," Journal of Techniques, vol. 5, no. 3, pp. 116–128, Sep. 2023, doi: 10.51173/jt.v5i3.1134.

[7] Noor Sattar Noor, Dalal Abdulmohsin Hammood, and Ali Al-Naji, "Applying TTIED-CMYK Algorithm in Wireless Sensor Networks Based on Raspberry pi and DHT-11," Journal of Techniques, vol. 4, no. 3, pp. 1–7, Sep. 2022, doi: 10.51173/jt.v4i3.593.

[8] S. M. Kadhim, J. Koh Siaw Paw, Y. C. Tak, and S. Ameen, "Deep Learning Models for Biometric Recognition based on Face, Finger vein, Fingerprint, and Iris: A Survey," Journal of Smart Internet of Things, vol. 2024, no. 1, pp. 117–157, Jun. 2024, doi: 10.2478/jsiot-2024-0007.

[9] M. Mohy-eddine, A. Guezzaz, S. Benkirane, and M. Azrour, "An efficient network intrusion detection model for IoT security using K-NN classifier and feature selection," Multimedia Tools and Applications, vol. 82, no. 15, pp. 23615– 23633, Feb. 2023.

[10] S. Hosseini and S. R. Sardo, "Network intrusion detection based on deep learning method in internet of thing," Journal of Reliable Intelligent Environments, vol. 9, no. 2, pp. 147–159, Feb. 2022.

[11] N. F. Syed, M. Ge, and Z. Baig, "Fog-cloud based intrusion detection system using Recurrent Neural Networks and feature selection for IoT networks," Computer Networks, vol. 225, 2023.

[12] C. Chen, Y. Gao, S. Huang, and X. Yan, "Avoid attacks: A Federated Data Sanitization Defense in IoMT Systems," IEEE INFOCOM 2023 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), May 2023.

[13] A. Binbusayyis, H. Alaskar, T. Vaiyapuri, and M. Dinesh, "An investigation and comparison of machine learning approaches for intrusion detection in IoMT network," The Journal of Supercomputing, vol. 78, no. 15, pp. 17403–17422, May 2022.

[14] Y. Manchala, J. Nayak, and H. S. Behera, "Detection of Malicious Traffic in IoMT Environment Using Intelligent XGboost Approach," 2022 OPJU International Technology Conference on Emerging Technologies for Sustainable Development (OTCON), Feb. 2023.

[15] J. Jithish, S. Sankaran, and K. Achuthan, "A Hybrid Machine Learning Approach for Intrusion Detection in Cyber-Physical Manufacturing Systems," Intelligent Security Solutions for Cyber-Physical Systems, pp. 156–168, Mar. 2024.

[16] E. M. Maseno and Z. Wang, "Hybrid wrapper feature selection method based on genetic algorithm and extreme learning machine for intrusion detection," Journal of Big Data, vol. 11, no. 1, Feb. 2024.

[17] H. Mohammed Fadhil, Z. O. Dawood, and A. Al Mhdawi, "Enhancing Intrusion Detection Systems Using Metaheuristic Algorithms," Diyala Journal of Engineering Sciences, pp. 15–31, Sep. 2024, doi: 10.24237/djes.2024.17302.

[18] "NSL-KDD | Datasets | Research | Canadian Institute for Cybersecurity | UNB," www.unb.ca. https://www.unb.ca/cic/datasets/nsl.html.

[19] smart gid dataset https://www.kaggle.com/datasets/hussainsheikh03/smart-grid-intrusion-detection-dataset.

[20] unsw_nb 2015 dataset https://research.unsw.edu.au/projects/unsw-nb15-dataset

[21] A. Wang, W. Wang, H. Zhou, and J. Zhang, "Network Intrusion Detection Algorithm Combined with Group Convolution Network and Snapshot Ensemble," Symmetry, vol. 13, no. 10, pp. 1814, Sep. 2021.