Journal of Information Systems Engineering and Management
2025, 10(19s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

A Novel Framework of Anomaly-Based Network Intrusion Detection using Hybrid CNN, Bi-LSTM Deep Learning Techniques

Srinivas Akkepalli [1], Sagar.K [2]
[1]*Asst prorofessor, Sree Chaitanya College of Engineering, Karimnagar, Telangana, India*
[2]*Principal,* Geethanjali College of Engineering and Technology, *Hyderabad, Telangana, India*

ARTICLEINFO

Received: 24 Dec 2024

Revised: 31 Jan 2025

Accepted: 14 Feb 2025

ABSTRACT

A Novel Framework of Anomaly-based Network Intrusion Detection system using hybrid CNN,Bi-LSTM Deep learning techniques with the aim of anomaly detection, In recent years, deep learning (DL) has become increasingly important in the field of cyber security. Deep learning Algorithms efficient to detect vulnerabilities in network traffic.Objective are based on literature survey provides the various anomaly based techniques Such as NIDS,SIDS, researches are presented. [proposed CNN based BLISTM model]stand out, providing a solid basis for understanding the context of the investigation and verified results with slandered existing systems and studies.The methodology adopted for this research comprises [a hybrid CNN-based BLISTM model with Adam optimizer was used] and [the well-known NSL KDD data set was used to validate the proposed modelThe results obtained revealed that efficacy of the suggested CNN-Bi-LSTM IDS has been assessed for the NSL-KDD dataset. Imbalanced data fed into CNN-Bi-LSTM accuracy achieved 98 % recall 98% and precision 99 %, F1-score98 %, After balanced data and hyper parameter tuning of CNN-Bi-LSTM classifier, Exceptional accuracy was demonstrated by the binary classification results, which included a 99.12% accuracy for the NSL KDD dataset with the precision of 99.0%, recall of 99.26%, and F1-score of 98.11%.. Conclusions are a novel frame work enhanced accuracy in detection of anomalies in network traffic in the field of[Network security]. These implications could encompass list impacted are such as Medical and Banking, Ecommerce sites.This study contributes to the literature by hybrid CNN based BLISTM generates efficacious results. The relevance and value of this research are evidenced by comparison of generated results with existing literature results. In future work it would applied for various different datasets .

Keywords: CNN,BLISTM,ADAM optimizer,NSL-KDD Dataset

INTRODUCTION

Traditional machine learning (ML) based anomaly detection systems primarily use manually extracted features from network traffic to classify and detect anomalies. These techniques still have a high false positive rate, which makes it difficult to detect any new or unidentified (0-day) attacks and drastically decreases the efficacy of in-time detection. It also adds a significant amount of manual scrutiny work to the process. In contrast, Deep Learning systems have demonstrated the ability to identify new characteristics and attack patterns in this dynamic environment, allowing them to uncover new attacks in addition to analysing the features that were manually extracted. These systems can also extract features from the original traffic[1]. In the suggested model, a dataset's spatial and high-level features are extracted and learned by CNN, while the long-range temporal features are learned by Bi-LSTM layers. These two networks are then combined to create a hybrid model that predicts attacks. The renowned NSL-KDD network intrusion dataset is then employed to assess the model, and the results show that this offers a lower false positive rate and better detection ability (better detection rate and validation accuracy. Below I describe the dataset and the, necessary preprocessing steps to train this model. Along with describing the hyperparameters and optimization techniques employed, I also discuss the model's architecture. Finally, I demonstrate this model's assessment based on various metrics. To determine the efficacy of the proposed CNN and Bi-LSTM-based DL model, as a performance baseline, a contrast with an ML-based approach is also conducted.

OBJECTIVES

Anomaly-based intrusion detection presents several additional challenges that can impact its effectiveness and efficiency. First one class imbalance in training data in the context of ML, the training data often contains a significant imbalance between normal and anomalous instances, with normal instances vastly outnumbering anomalies. This imbalance can lead to biased models that are less effective at detecting rare but critical anomalies. Second context awareness anomaly detection systems often lack contextual understanding. For instance, an activity that is normal in one context of a specific time of day,might be anomalous in another. Incorporating context-awareness into anomaly detection models is complex but necessary for reducing false positives and improving accuracy. Third High False Positive Rate one of the most significant challenges in anomaly-based IDS is the high rate of false positives. Since this approach identifies deviations from normal behavior, any legitimate but unusual activity can be flagged as an anomaly, leading to numerous false alarms. Fourth defining normal behavior accurately defining normal behavior in a network or system environment is difficult. Normal behavior can vary widely over time due to changes in user behavior, network traffic patterns, and system configurations. Establishing a precise and adaptive baseline is crucial but challenging. Fifth evolution of threats cyber threats continuously evolve, and attackers often change their tactics to evade detection. Anomaly-based systems need to be constantly updated and refined to recognize new patterns of normal behavior and new types of anomalies.

This paper aims to give a solution for two issues the design and develop a preprocessing method to handle imbalanced data using a Minority oversampling algorithm. Second design and develop a Hybrid DL model, CNN-based BiLSTM for anomaly detection.

METHODS

A DataSet and Features

The most common sources of datasets in this domain are auditing events from applications or network traffic from different network devices. To assess the model, we take a look at the well-known NSL-KDD dataset: The NSL-KDD dataset, which includes both attack and benign vectors, is an improved version of the original KDD99 dataset produced by the CICS ("Canadian Institute for Cyber Security"). 42 attributes constitute the NSL-KDD dataset, and the class attribute includes five attacks that are divided into four groups: root-to-local (R2L), probing attacks, DoS ("Denial-Of-Service"), and U2R ("User-To-Root"). For multi-category attack prediction, 5 classes including the Normal category have been employed in this dataset.

B Data Preprocessing

One-hot encoding of categorical features and normalization of numeric features are the methods used to preprocess the datasets. 1. One hot encoding: To enhance model training, categorical columns like "protocol," "state," and "service" are one hot encoded.

C Normalization

A rescaling of the input data is performed between 0 and 1 using the feature scaling technique known as min-max normalization. Improved regularization effect and improved gradient descent convergence are the driving forces behind this use.

D. Proposed Methodology

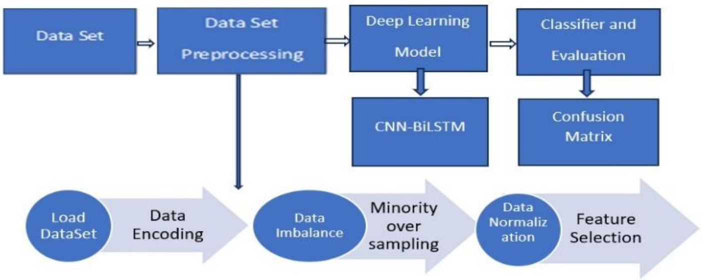

Fig .1. Process flow diagram

The suggested model was trained with an Anaconda Jupiter notebook and is implemented with the help of the Keras, TensorFlow backend, and Scikit-learn packages. The suggested model combines several layers of the Bidirectional LSTM (Bi-LSTM) network with "a 1-D CNN (1-Dimensional Convolutional Neural Network). We employ 1D convolution because the network packet is expressed in a 1D format. Lever-aging the 1-D CNN layer aims to improve coarse-grained feature extraction at the start of the network and enable the network to extract features from input data. Three characteristics are combined: parameter sharing, spatial invariance, and max pooling layers. While spatial invariance makes it possible to identify feature correlation more accurately, parameter sharing allows for the restriction of the number of parameters in the initial layers, resulting in feature extraction requiring less computational re-sources. For faster convergence, the activation function for this layer is the Rectified Linear Unit (ReLU). In order to minimize the impact of covariance shifts from layer to layer throughout training, stop slower learning, and enable more robust weight fluctuations in the network, batch normalization is used between the layers that are intermediate. Subsequently, Reshape Layers are employed to modify the output of the preceding layer for incorporation into the ensuing layers. The spatial relations of the data are frequently displayed in the eature map that CNN generates. When it comes to data with long-range dependencies, CNN performs poorly. As opposed to this, RNN can retrieve temporal structures from the input data. Every iteration of the model doubles the kernel size due to the arrangement of its 2 Bi-LSTM layers (64 & 128 units). To gain more insight into the relationship between features learned by the 1st 1-D CNN layer and long-range time-dependent features learned by the LSTM layers that follow, an attempt was made to replicate the usage of coarse-grained to fine-grained learning.

To improve performance and shorten training times, there are two layers in between every Bi-LSTM layer: a Max Pooling layer that eliminates the least important features and a Batch Normalization layer that normalizes the output data of the preceding intermediate layer. Since the granularity of the output varies between LSTM layers, a reshape layer is added to reshape the input data for the next LSTM block. The fully connected layer" functions as the output layer in the model and has a sigmoid activation function for binary classification. To improve generalization during model inference, a regularization technique called Adam dropout is used, and its rate value is 0.6. Binary cross entropy is utilized for binary classification, and categorical cross entropy is employed when the labels are one hot encoded [11]. There are 32 batch sizes set.

E.Proposed BiLSTM Methodology

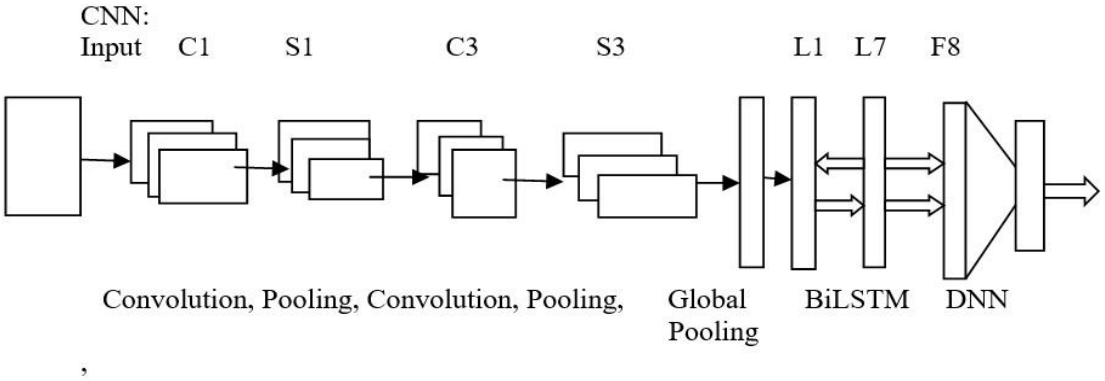

Fig.2. CNN Model

Input data size(I), Feature detector(K), pooling (M), Stride(S), Flatten(F)

OutputI.K= ([I-K]/S)+1 (1)

Pooling= floor(I/2) (2)

$$z^{l+1}(i,j)=\sum_{k=1}^{c}\sum_{x=1}^{f}\sum_{y=1}^{f}\left[Z_K^1(s*i+x,s*j+y)*\omega_k^{l+1}(x,y)\right]+b \quad (3)$$

The convolution layer contains an activation function (formula 12) that assists in the expression of complex features.K is the number of channels in the characteristic graph and A represents the output vector of the Z vectorthrough the activation function. We used sigmoid and ReLU, respectively, after two convolution layers.

$A^1_{I,J,K} = f(Z^1_{I,J,K})$ (4)

After feature extraction in the convolution layer, the output image is transferred to the pooling layer for feature.Selection and information filtering. The pooling layer contains a preset pooling function that replaces the result of a single point in the feature map with the feature graph statistics of its adjacent region.

The pooling layer is calculated by formula 13, where p is the pre-specified parameter.

$A^l_k(i,j)\left[\sum_{x=1}^{n}\sum_{y=1}^{n} A^l_k(s*i+x, s*j+y)^p\right]$^1/p (5)

We also used a back-propagation algorithm to adjust the model parameters. In the weight adjustment algorithm. (Formula 6), δ is delta error of loss function to the layer, and α is the learning rate

$\omega^l = \omega^l - \alpha\sum \delta^l * A^{l-1}$ (6)

We used the categorical cross-entropy algorithm in the loss function. In order to re-duce training time and enhance the gradient descent accuracy. we used the Adams optimization and after two convolution and pooling operations, we extracted the entire traffic image into a smaller feature block, which represents the feature information of the whole traffic packet. The block can then be fed into the RNN system as an input to the RNN layer.

F. operation of BILSTM

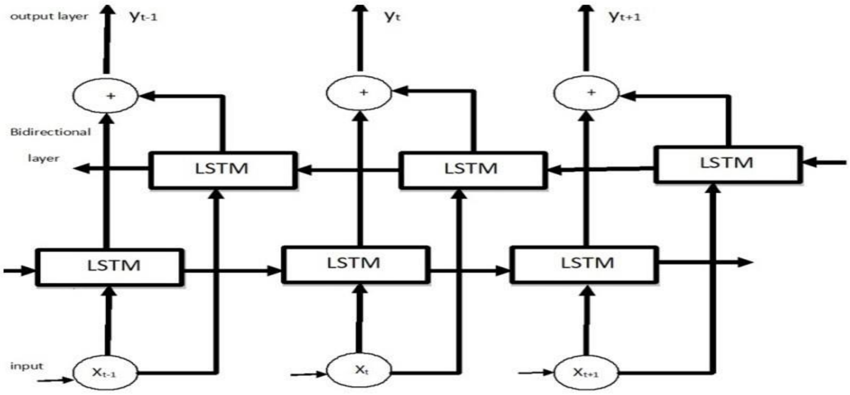

Fig.3. Bi-LSTM

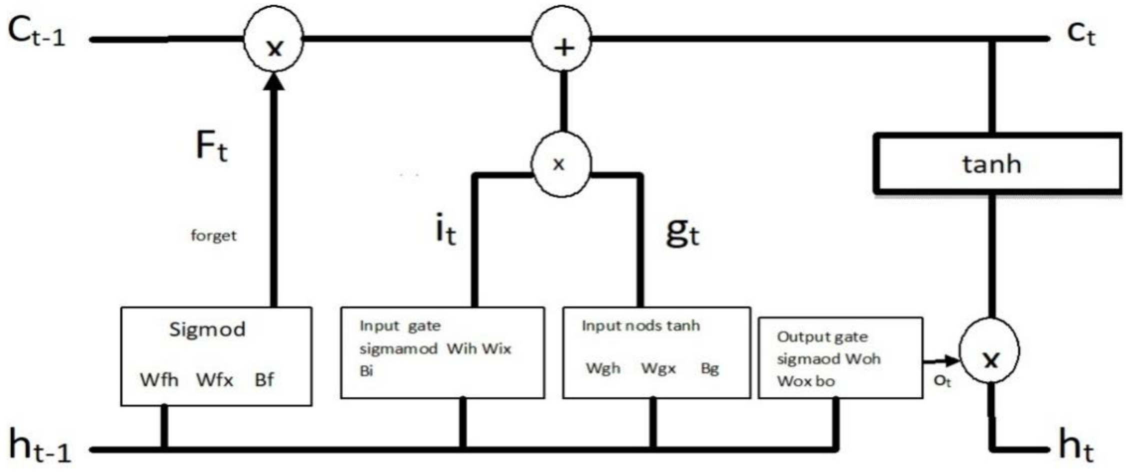

Fig.4. Single unit LSM

ft =[(Wfh*ht-1)+(Wfx*xt)+bf] (7)

it =[(Wih *ht-1)+(Wgx *xt)+bi (8)

gt =tanh[(wgh*ht-1)+(Wgx *xt)+bg] (9)

ot =[(Woh *ht-1)+(wox * xt)+bo] (10)

ct= ct-1x ft + itxgt (11)

ht= tanh(ct)xot (12)

RESULTS

With batch norm and dropouts applied at every layer, the model is composed of multiple blocks of 1DConvolutional and LSTM layers. I experimented with this mod-el. A fully connected layer with sigmoid activation, a global average pooling layer, and a 1D-convolutional layer constitute the final layers. It is as promising as it is effective in terms of detection rate and accuracy" in particular, the detection rates and accuracy in relation to the other models. Anaconda Jupiter notebook was used for the model's training, and Tensor Flow and Keras' Scikit-learn metrics Python library were used for the evaluation. The model's performance in binary classification determines whether the data indicates an attack or normal traffic. This involves mean result, as estimated employing the standard error, is anticipated to be a more accurate estimation of the true unidentified underlying average performance of the model on the dataset. Three main metrics are used to examine the performance of the models on the NSL-KDD dataset: FPR, Accuracy, and Detection Rate. The model's accuracy quantifies its capacity for predicting binary classes. The rate of recall or detection (TP / TP + FN). evaluates the model's accuracy in predicting attacks. The fraction of normal records that are classified as attacks is known as precision, or FPR (FP / FP+TN). The F-1 score yields a single score for the harmonic mean of recall and precision in a single number.

Precision =TP /(TP+ FP)

Recall =TP /(TP+ FN)

Accuracy =TP+ TN /(TP+ FP+ TN +FN)

F-measure=2×Precision×Recall/(Precision+ Recall)

Table 1 . Comparison Of Results

Detection Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
(2018) Hera Naseer et al. [1] DCNN, LSTM, Decision Tree	85.89	94	98	93.05
(2019) Vinay Kumar et al. [2] DNN based stacking of MLP	97	96	97	96.49
(2020) BLSTM Multiple CNN layers [3]	84.25	84	85.5	85
(2021) Deep Stacking Network for Intrusion detection [4]	90.41	96.65	78.82	86.83
(2022) LSTM, GRU [5]	94.2	97	98	97.49
(2023) RNN, LSTM [6]	95.0	94.7	99	96.8
Proposed Model	99.28	99	99.26	99.18

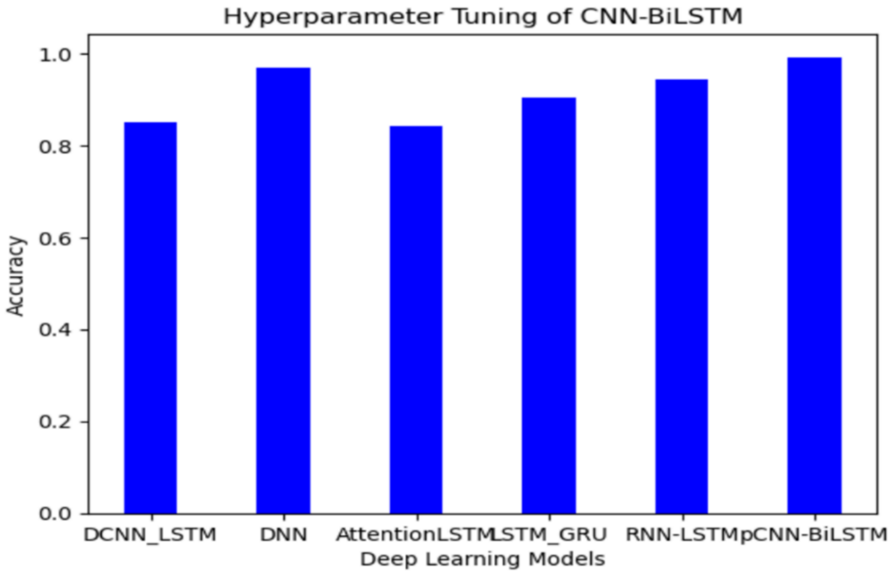

Fig 5 .Comparison of results with existing Anomaly based Network Intrusion detection Systems

DISCUSSION

This research suggests a novel method for intrusion detection in network security that combines statistical analysis and DL techniques. The model shows notable advancements in computer network intrusion detection. Using standard measurement systems, the efficacy of the suggested CNN-Bi-LSTM IDS has been assessed for the NSL-KDD dataset. Imbalanced data fed into CNN-Bi-LSTM accuracy achieved 98 % recall 98% and precision 99 %, F1-score98 %, After balanced data and hyper parameter tuning of CNN-Bi-LSTM classifier, Exceptional accuracy was demonstrated by the binary classification results, which included a 99.28% accuracy for the NSL KDD dataset with the precision of 99.0%, recall of 99.26%, and F1-score of 98.18%. Further research can be conducted by using DL classifiers to identify intrusions in more real-time and online datasets.

REFRENCES

[1] heraz naseer1, yasir saleem2,Enhanced Network Anomaly Detection Based on Deep Neural Networks, ,vol 6. 2018. doi 10.1109/access.2018.286303, IEEE Access

[2] vinayakumar r1 ,mamoun alazab2 , (senior member, ieee), somankppraba-haranpoornachandran, Deep Learning Approach for Intelligent Intrusion Detection System, , 10.1109/ACCESS.2019.2895334, IEEE Access.

[3] TONGTONG SU, HUAZHI SUN , JINQI ZHU ,SHENG WANG , AND YABO LI, BAT: Deep Learning Methods on Network Intrusion Detection using NSL-KDD dataset., : DOI 10.1109/ACCESS.2020.2972627, IEEE Access..

[4] Yifan Tang, Lize Gu * and Leiting Wang, Deep Stacking Network for Intrusion Detection, Sensors-2021, MDPI, https://doi.org/10.3390./s22010025.

[5] EZAT AHMADZADEH, HYUNIL KIM, ONGEE JEONG, NAMKI KIM, AND INKYU MOON, (Member, IEEE),A Deep Bidirectional LSTM-GRU Network Model for Automated Cipher-text Classification. date of publication Janu-ary 4, 2022, date of current version January 10, 2022. Digital Object Identifier 10.1109/IEEE ACCESS.2022.3140342.

[6] Mariam Ibrahim, Rube El hafiz,(Modelling an intrusion detection using recurrent neural networks Journal of Engineering Research March 2023, Elsevier.

[7] Peilun Wu and Hui Guo, “LuNet: A Deep Neural Network for Network Intrusion Detection” 10.1109/ SSCI44817.2019.9003126, Dec 2019.

[8] Puneet Misra1, Arun Singh Yadav, International Improving the Classification Accuracy using Recursive Feature Elimination with Cross-Validation, Journal on Emerging Technologies,May 2020. 11(3): 659665(2020), ISSN No: (Online): 2249-3255.

[9] Thomas Rincy N and Roopam Gupta,Design and Development of an Efficient Network Intrusion Detection System Us-ing Machine Learning Techniques Wire-less Communications and Mobile Computing Volume 2021, Article ID 9974270, 35 pages https://doi.org/10.1155/2021/9974270.

[10] ShadiAljawarneh, MontherAldwairi,Muneer Bani Yassein, Anomaly-based intrusion detection system through feature selection analysis and building ybrid efficient model, ournal of Computational Science(2017),Elsevier.

[11] Zhang, X.; Wen, S.; Yan, L.; Feng, J.; Xia, Y. A Hybrid-Convolution Spatial–Temporal Recurrent Network For Traffic Flow Prediction. Comput. J. 2022, bxac171

[12] Han, Z.; Yang, Y.; Wang, W.; Zhou, L.; Gadekallu, T.R.; Alazab, M.; Gope, P.; Su, C. RSSI map-based trajectory design for UGV against malicious radio source: A reinforcement learning approach. IEEE Trans. Intell. Transp. Syst. 2022, 24, 4641–4650.

[13] Schufrin, M.; Lücke-Tieke, H.; Kohlhammer, J. Visual Firewall Log Analysis-At the Border Between Analytical and Appealing. In Proceedings of the 2022 IEEE Symposium on Visualization for Cyber Security (VizSec), Oklahoma City, OK, USA, 19 October 2022; pp. 1–11.

[14] Xue, B.; Warkentin, M.; Mutchler, L.A.; Balozian, P. Self-efficacy in information security: A replication study. J.Comput. Inf. Syst. 2023, 63, 1–10. [CrossRef].

[15] Yu,JLu,L.; Chen, Y.; Zhu, Y.; Kong, L. Anindirecteavesdroppingattackofkey-strokesontouchscreenthroughacousticsensing. IEEE Trans. Mob. Comput. 2019, 20, 337–351. [CrossRef].

[16] Alsharif, M.; Mishra, S.; AlShehri, M. Impact of Human Vulnerabilities on Cy-bersecurity. Comput. Syst. Sci. Eng. 2022, 40, 1153–1166. [CrossRef].

[17] Margossian, H.; Sayed, M.A.; Fawaz, W.; Nakad, Z. Partial grid false data injec-tion attacks against state estimation. Int. J. Electr. Power Energy Syst. 2019, 110, 623–629. [CrossRef].

[18] Wahab, O.A.; Bentahar, J.; Otrok, H.; Mourad, A. Resource-aware detection and defense system against multi-type attacks in the cloud: Repeated bayesian-stackelberg game. IEEE Trans. Dependable Secur. Comput. 2019, 18, 605–622. [CrossRef].

[19] Wahab, O.A.; Bentahar, J.; Otrok, H.; Mourad, A. Optimal load distribution for the detection of VM-based DDoS attacks in the cloud. IEEE Trans. Serv. Com-put. 2017, 13, 114–129. [CrossRef].

[20] Kavitha, C.; Gadekallu, T.R.; Kavin, B.P.; Lai, W.C. Filter-Based Ensemble Fea-ture Selection and Deep Learning Model for Intrusion Detection in Cloud Com-puting. Electronics 2023, 12, 556. [CrossRef].

[21] Shaikh, S.; Rupa, C.; Srivastava, G.; Gadekallu, T.R. Botnet Attack Intrusion De-tection In IoT Enabled Automated Guided Vehicles. In Proceedings of the 2022 IEEE International Conference on Big Data (Big Data), Osaka, Japan, 17–20 December 2022; pp. 6332–6336.

[22] Dbouk, T.; Mourad, A.; Otrok, H.; Tout, H.; Talhi, C. A novel ad-hoc mobile edge cloud offering security services through intelligent resource-aware offload-ing. IEEE Trans. Netw. Serv. Manag. 2019, 16, 1665–1680. [CrossRef].

[23] Rani, S.; Babbar, H.; Srivastava, G.; Gadekallu, T.R.; Dhiman, G. Security Framework for Internet of Things based Software Defined Networks using Blockchain. IEEE Internet Things J. 2022, 10, 6074–6081. [CrossRef].

[24] Kong, H.; Lu, L.; Yu, J.; Chen, Y.; Tang, F. Continuous authentication through finger gesture interaction for smart homes using WiFi. IEEE Trans. Mob. Comput. 2020, 20, 3148–3162. [CrossRef].

[25] Nagasree, Y.; Rupa, C.; Akshitha, P.; Srivastava, G.; Gadekallu, T.R.; Laksh-manna, K. Preserving privacy of classified authentic satellite lane imagery using proxy re-encryption and UAV technologies. Drones 2023, 7, [CrossRef].

[26] Shamseddine, H.; Nizam, J.; Hammoud, A.; Mourad, A.; Otrok, H.; Harmanani, H.; Dziong, Z. A novelfederatedfogarchitecture embedding intelligent formation. IEEE Netw. 2020, 35, 198–204. [CrossRef].

[27] Srivastava, G.; K, D.R.R.; Yenduri, G.; Hegde, P.; Gadekallu, T.R.; Maddikunta, P.K.R.; Bhattacharya, S. Federated Learning Enabled Edge Computing Security for Internet of Medical Things: Concepts, Challenges and Open Issues. In Securi-ty and Risk Analysis for Intelligent Edge Computing; Springer: Berlin/Heidelberg, Germany, 2023; pp. 67–89.

[28] AbdulRahman, S.; Tout, H.; Mourad, A.; Talhi, C. FedMCCS: Multicriteria client selection model for optimal IoT federated learning. IEEE Internet Things J. 2020, 8, 4723–4735. [CrossRef].

[29] AbdulRahman, S.; Tout, H.; Mourad, A.; Talhi, C. FedMCCS: Multicriteria client selection model for optimal IoT federated learning. IEEE Internet ThiDPengfei Sun,1engju Liu, Qi Li,Chenxi Liu, Xian-glingLu,RuochenHao,andJinpengChenL-IDS: Extracting Features Using CNN-LSTM Hybrid Network for Intrusion Detection System, Security and Communi-cation Networks Volume 2020, Article ID 8890306, 11 pages https://doi.org/10.1155/2020/8890306ngs J. 2020, 8, 4723–4735. [CrossRef].

[30] Sarkar, S.; Agrawal, S.; Gadekallu, T.R.; Mahmud, M.; Brown, D.J. Privacy-Preserving Federated Learning for Pneumonia Diagnosis. In International Con-ference on Neural Information Processing, Proceedings of the 29th International Conference, ICONIP 2022, Virtual Event, 22–26 November 2022; Proceedings, Part VII; Springer: Berlin/Heidelberg, Germany, 2023; pp. 345–356.

[31] Aysha Bibi , Gabriel Avelino Sampedro, Ahmad Almadhor, Abdul Rehman Javed and Tai-hoon Kim ,Hypertuned Lightweight and Scalable LSTM Model for Hybrid Network Intrusion Detection.

[32] TONGTONG SU1,a HUAZHI SUN , JINQI ZHU ,SHENG WANG , AND YABO LI, BAT: Deep Learning Methods on Network Intrusion Detection using NSL-KDD dataset.

[33] Mariam Ibrahim,RubaElhafiz ,Modeling an intrusion detection using recurrent neural networks,Journal of Engineering Research Volume 11, Issue 1, March 2023, 100013.