Download   Print   Save to PDF   Save to list   Create bibliography

Multi-scale Memory Residual Network Based Deep Learning Model for Network Traffic Anomaly Detection

Jayakrishna M.[a] ✉ ; Selvakumar V.[b]; Kumar, Atul[c]; Dilip, Salunke Mangesh[d]; Maaliw, Renato R.[e] ✉

Save all to author list

[a] Mechanical engineering, Sri Sivani Engineering College, Srikakulam, India

[b] Department of Maths and Statistics, Bhavan's Vivekananda College of Science, Humanities and Commerce, Telangana, Hyderabad, India

[c] Dr. D. Y. Patil B-School, Pune, India

[d] Department of Computer Engineering, GHRCEM, Pune, India

[e] College of Engineering, Southern Luzon State University, Lucban, Quezon, Philippines

Hide additional affiliations

Full text options   Export

Abstract

Author keywords

Indexed keywords

SciVal Topics

Metrics

Abstract

Models for detecting network traffic anomalies based on deep learning usually exhibit weak generalization, confined representative capacity, and low real-world adaption. In light of this, a multi-scale memory residual network-based model for identifying network traffic anomalies is proposed. The

Cited by 0 documents

Related documents

Deep Learning-Based Intrusion Detection Model for Network Security

Pande, S.D. , Lanke, G.R. , Soni, M.

(2023) Lecture Notes in Networks and Systems

Machine Learning with Variational AutoEncoder for Imbalanced Datasets in Intrusion Detection

Lin, Y.-D.   Liu, Z.-Q.   Hwang, R.-H.
,   ,

(2022) IEEE Access

Analysis Method of Abnormal Traffic of Teaching Network in Higher Vocational Massive Open Online Course Based on Deep Convolutional Neural Network

Chen, H.   Zou, J.
,

(2023) Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)

View all related documents based on references

Find more related documents in Scopus based on:
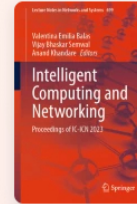
Authors >   Keywords >

Home > Intelligent Computing and Networking > Book Chapter

Multi-scale Memory Residual Network Based Deep Learning Model for Network Traffic Anomaly Detection

Book chapter | First Online: 05 August 2023

pp 475–482 | Cite this book chapter

Intelligent Computing and Networking

(IC-ICN 2023)

M. Jayakrishna ✉, V. Selvakumar, Atul Kumar, Salunke Mangesh Dilip & Renato R. Maaliw

🔊 330 Accesses

Abstract

Models for detecting network traffic anomalies based on deep learning usually exhibit weak generalization, confined representative capacity, and low real-world adaption. In light of this, a multi-scale memory residual network-based model for identifying network traffic anomalies is proposed. The distribution analysis of the three-dimensional feature space illustrates the efficiency of the network traffic data preprocessing technique. The deep learning algorithm enhances the model's capacity to represent data by coupling multi-scale 1DCNN-LSTM networks. The realization of the residual network is shown using the residual network notion as a foundation. Deep feature extraction accelerates model convergence to detect network traffic anomalies accurately and effectively while preventing gradient disappearance, explosion, over fitting, and network damage. The experimental findings show how the multi-scale 1DCNN-LSTM network can improve the model's representational competence and generalization ability. Performance indicators for the model in this study are also superior to those for other deep learning models.

Access this chapter

Log in via an institution →

∧ Chapter   EUR 29.95
                                Price includes VAT (Philippines)

- Available as PDF
- Read on any device
- Instant download
- Own it forever

Buy Chapter →

Copyright information

About this paper

Publish with us

References

1. Hwang R-H, Peng M-C, Huang C-W, Lin P-C, Nguyen V-L (2020) An unsupervised deep learning model for early network traffic anomaly detection. IEEE Access 8:30387–30399. https://doi.org/10.1109/ACCESS.2020.2973023

Multi-scale Memory Residual Network Based Deep Learning Model for Network Traffic Anomaly Detection

M. Jayakrishna, V. Selvakumar, Atul Kumar, Salunke Mangesh Dilip, and Renato R. Maaliw

Abstract Models for detecting network traffic anomalies based on deep learning usually exhibit weak generalization, confined representative capacity, and low real-world adaption. In light of this, a multi-scale memory residual network-based model for identifying network traffic anomalies is proposed. The distribution analysis of the three-dimensional feature space illustrates the efficiency of the network traffic data preprocessing technique. The deep learning algorithm enhances the model's capacity to represent data by coupling multi-scale 1DCNN-LSTM networks. The realization of the residual network is shown using the residual network notion as a foundation. Deep feature extraction accelerates model convergence to detect network traffic anomalies accurately and effectively while preventing gradient disappearance, explosion, over fitting, and network damage. The experimental findings show how the multi-scale 1DCNN-LSTM network can improve the model's representational competence and generalization ability. Performance indicators for the model in this study are also superior to those for other deep learning models.

Keywords Deep learning · Feature extraction · Network traffic anomaly detection · Network degradation · One-Hot encoding · CNN · LSTM

M. Jayakrishna (✉)
Mechanical engineering, Sri Sivani Engineering College, Srikakulam, India
e-mail: jayakrishnamakka555@gmail.com

V. Selvakumar
Department of Maths and Statistics, Bhavan's Vivekananda College of Science, Humanities and Commerce, Hyderabad, Telangana, India

A. Kumar
Dr. D. Y. Patil B-School, Pune, India

S. M. Dilip
Department of Computer Engineering, GHRCEM, Pune, India

R. R. Maaliw
College of Engineering, Southern Luzon State University, Lucban, Quezon, Philippines
e-mail: rmaaliw@slsu.edu.ph

1 Introduction

With the development of network technology and the expansion of network scale, network traffic is increasing exponentially, and network security threats and risks are becoming more and more prominent. Intrusion Detection System (IDS) [1] is a network security monitoring system. Network traffic is one of the main network states. When network intrusions occur, network traffic anomalies usually occur. Therefore, network traffic anomaly detection is the current network Research focus of the Network Intrusion Detection System (NIDS).

However, the continuous change of network attack patterns has increased the difficulty of network traffic anomaly detection [2]. Based on the empowerment effect of artificial intelligence, cyberspace security is facing new risks, including network attacks becoming more and more intelligent, and large-scale attacks becoming more and more difficult. Network attacks are becoming more and more frequent, the concealment of network attacks is getting higher and higher, the game-fighting nature of network attacks is becoming stronger and stronger, and important data is becoming easier to steal, etc. [3]. Maintaining network security is a process of attack and defense games. Network traffic Anomaly detection, as a prerequisite for ensuring network security, has received more and more attention because it can identify unknown network attacks.

The representational capacity is constrained and the false alarm rate is high [4]. It is challenging for classical machine learning to accomplish the goal of analysis and prediction due to the increase in huge data on the network [5–7], the improvement of network capacity [8–11], the complexity of data [12–15], and the diversity of features [4]. Large-scale network traffic data can be processed successfully using the deep learning approach [16–19]. Deep learning [20–22] offers better representation performance when compared to conventional machine learning techniques, which can significantly increase the effectiveness [23, 24] and precision of network traffic anomaly identification [25, 26]. The identification of network traffic anomalies is a powerful tool for thwarting contemporary network attacks. The network traffic anomaly detection model is the main topic of this study, which also suggests a multi-scale memory residual network-based network traffic anomaly detection model.

2 MSMRNet Based IDS Model

Since standard recurrent neural networks have an issue with gradient disappearance, the long-short-term memory neural network (LSTM) has been presented as a solution (RNN).Its basic unit is a structure containing multiple groups of neurons, called a cell (cell), as shown in Fig. 1.

Combining the idea of the residual network and long-term short-term memory network, this paper proposes a network traffic anomaly detection model based on MSMRNet.
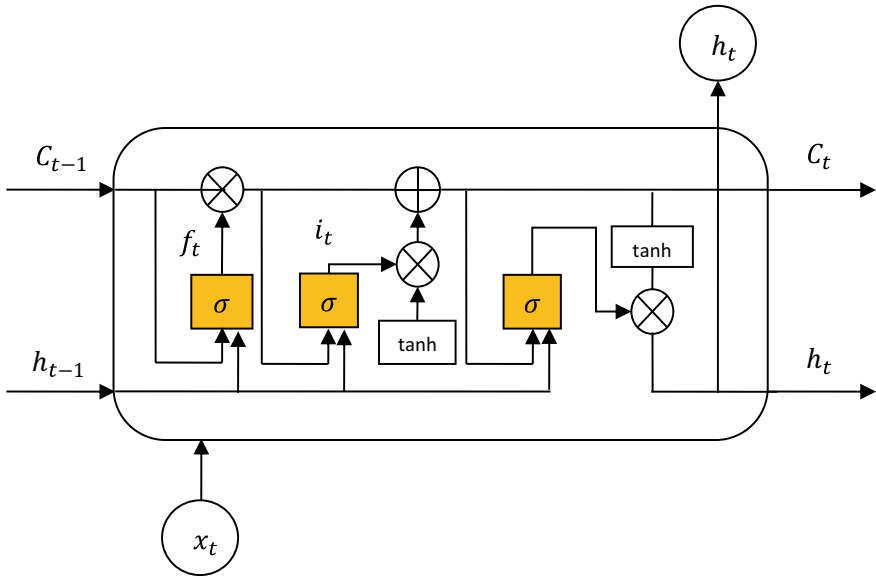
Fig. 1 Long short-term memory network

There are non-numeric data such as protocol type and service type in the network traffic data, and the machine learning model cannot handle non-numeric data, so the network traffic data needs to be processed numerically. At the same time, the network traffic data is different, and the characteristic attributes. There is a large difference in magnitude between [4], so it is necessary to carry out dimensionless processing on the network traffic data. The initial network traffic data $X_0 = \{x_1, x_2, \cdots, x_n\}$, where $f = |X_0|$ represents the characteristic dimension of the initial traffic data, and $n = |X|$ represents the characteristic dimension of the network traffic data after data preprocessing.

Input the preprocessed network traffic data X to MSMRNet for deep feature extraction, that is, the initial input of MSMRNet $X_1 = X$. MSMRNet is formed by stacking several multi-scale memory residual modules, and its l multi-scale memory residual the difference module takes Xl as input and generates output X_{l+1}, where the input X_l and output X_{l+1} have the same dimension. The multi-dimensional output Y_0 is one-dimensionalized by the Flatten layer to obtain the output Z_0. The local features are comprehensively processed through the fully connected layer [5] Obtain the output Z. Use the softmax function as the classifier to realize the network traffic classification, and obtain the network traffic classification result Y. Each element in Y represents the probability of each network traffic category, and the maximum probability category is the classification result. The calculation formula as shown in formula (1), where W_d and b_d represent the weight matrix and bias item respectively.

$$Y = softmax(Z) = softmax\left(W_d^T Z_0 + b_d\right)\dots\dots \qquad (1)$$

The implementation method of the network traffic anomaly detection model based on MSMRNet is shown in Algorithm 1.

Algorithm 1 Implementation method of network traffic anomaly detection model based on MSMRNet

Input: Training set X_{train}, Test set X_{test}, label set Y.
Output: network traffic classification result Y_D

Steps 1 Data preprocessing

1. Perform numerical processing on the training set X_{train} and the test set X_{test} to obtain $X_{\_train\_num}$ and $X_{\_test\_num}$
2. Perform dimensionless processing on the training set $X_{\_train\_num}$ and test set $X_{\_test\_num}$ to obtain X_{train} and X_{test}

Step 2 Build the model

3. Add several multi-scale memory residual modules
4. Add Flatten layer and fully connected layer, use softmax function as the classifier step 3 training model
5. Set the experimental hyper-parameters: optimizer, the number of single training samples batch_size, the learning rate , the number of training rounds epoch. Set the experimental verification set
6. While it did not reach the preset number of training rounds epochs do
7. While the training set is not empty
8. Take the mini-batch data set batch as the model input
9. Calculate the cross-entropy loss function, C represents the number of network traffic categories
10. Update model parameters using Adam optimizer
11. end while
12. Use the validation set to validate the model and perform parameter fine-tuning
13. end while

Step 4 Save the model

14. Save the Fine-Tuned Model

Step 5 Test the model

15. Load the saved model, test the model with the test set
16. return test set network traffic data classification results

3   Experiment

In order to obtain an effective data preprocessing method, this experiment uses the PCA two-dimensional visualization method to analyze the feature space distribution of the experimental data set after data preprocessing. The PCA two-dimensional
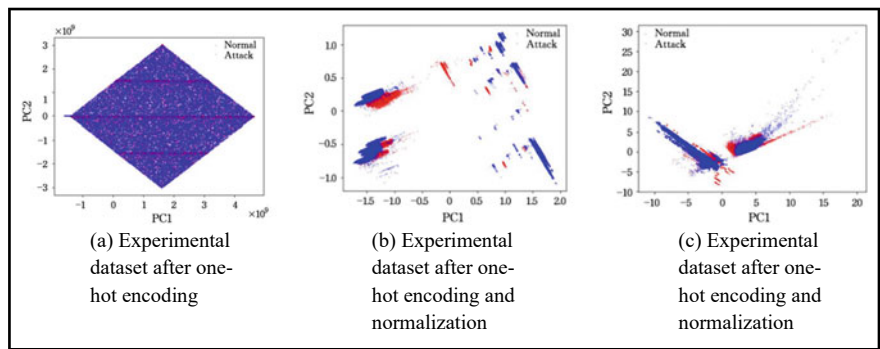
Fig. 2 Visualization of data set after preprocessing using PCA

visualization results of the experimental data set after one-hot encoding are shown in Fig. 2. As shown in (a), it can be seen that the normal flow data points overlap with the abnormal flow data points in a large area. The PCA two-dimensional visualization results of the experimental data set after one-hot encoding and standardization processing are shown in Fig. 2b, which can be It can be seen that the normal flow data points and the abnormal flow data points have been partially separated, but there are still some overlapping areas. The PCA two-dimensional visualization results of the experimental data set after one-hot encoding and normalization processing are shown in Fig. 2c, it can be seen that compared with the results after one-hot encoding and normalization, normal flow data points and abnormal flow data points have been effectively separated. Therefore, this experiment uses one-hot encoding to deal with.

3.1 Validity Verification Experiment

In order to verify the effectiveness of MSMRNet in solving the problem of network degradation, this paper constructs a multi-scale memory network (Multi-Scale Memory Network-work, MSMNet) of different depths and compares it with MSMRNet. The experimental model is as follows.

MSMNet-5: It is stacked by 5 multi-scale memory modules, including 20 training parameter layers, 10 non-training parameter layers and 1 fully connected layer.

MSMRNet-5: It is stacked by 5 multi-scale memory residual modules, including 20 training parameter layers, 10 non-training parameter layers and 1 fully connected layer.

MSMNet-10: It is stacked by 10 multi-scale memory modules, including 40 training parameter layers, 20 non-training parameter layers and 1 fully connected layer.
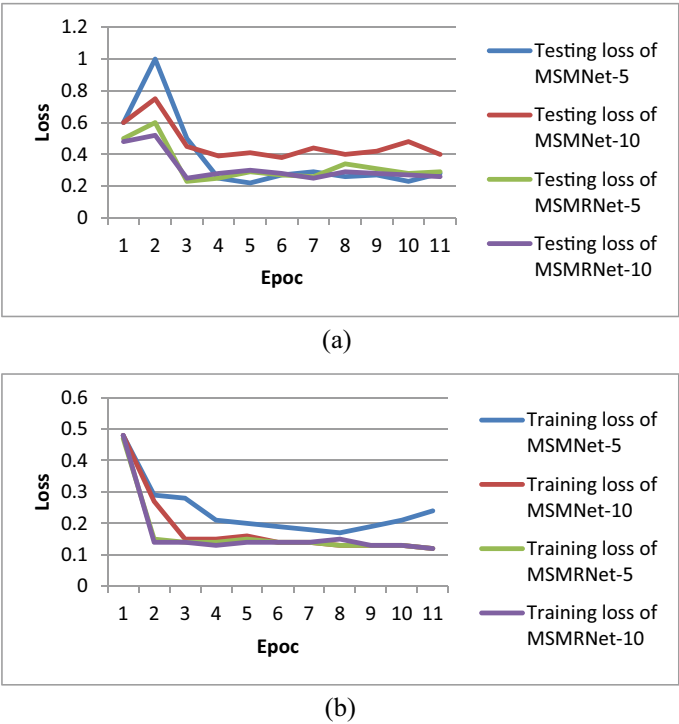
(a)

(b)

Fig. 3 a Comparison of loss rates between MSMNet and MSMRNet. **b** Comparison of loss rates between MSMNet and MSMRNet

MSMRNet-10: It is stacked by 10 multi-scale memory residual modules, including 40 training parameter layers, 20 non-training parameter layers and 1 fully connected layer.

Figure 3a, b show the comparison of the loss rate between MSMNet and MSMRNet during training and testing.

4 Conclusion

This work proposed a network traffic anomaly detection model based on the multi-scale memory residual network to address the issues of poor environmental adaptation, restricted representation ability, and weak generalization ability of the network traffic anomaly detection model based on deep learning. This essay is based on crediblity The effectiveness of the network traffic data preprocessing method is demonstrated by the analysis of the three-dimensional feature space distribution; the multi-scale one-dimensional convolution is combined with the long short-term memory network, and the model representation ability is enhanced through the deep

learning algorithm; based on the idea of the residual network, there is realization of in-depth feature extraction, while preventing gradient disappearance, gradient explosion, and over fitting. The results of data preprocessing visualization demonstrate that following one-hot encoding, the results of validity verification experiments and performance evaluation experiments show that adding identity mapping can accelerate model convergence, improve network traffic anomaly detection performance, and effectively solve the problem of network degradation; comparative experimental results show that normalization processing can effectively separate normal traffic and abnormal traffic data. Performance metrics have improved.

References

1. Hwang R-H, Peng M-C, Huang C-W, Lin P-C, Nguyen V-L (2020) An unsupervised deep learning model for early network traffic anomaly detection. IEEE Access 8:30387–30399. https://doi.org/10.1109/ACCESS.2020.2973023
2. Su T, Sun H, Zhu J, Wang S, Li Y (2020) BAT: deep learning methods on network intrusion detection using NSL-KDD dataset. IEEE Access 8:29575–29585. https://doi.org/10.1109/ACCESS.2020.2972627
3. Fernández Maimó L, Perales Gómez ÁL, García Clemente FJ, Gil Pérez M, Martínez Pérez G (2018) A self-adaptive deep learning-based system for anomaly detection in 5G networks. IEEE Access 6:7700–7712. https://doi.org/10.1109/ACCESS.2018.2803446
4. Wang W et al (2021) Anomaly detection of industrial control systems based on transfer learning. Tsinghua Sci Technol 26(6):821–832. https://doi.org/10.26599/TST.2020.9010041
5. Mezina A, Burget R, Travieso-González CM (2021) Network anomaly detection with temporal convolutional network and U-Net model. IEEE Access 9:143608–143622. https://doi.org/10.1109/ACCESS.2021.3121998
6. Han D et al (2021) Evaluating and improving adversarial robustness of machine learning-based network intrusion detectors. IEEE J Sel Areas Commun 39(8):2632–2647. https://doi.org/10.1109/JSAC.2021.3087242
7. Odiathevar M, Seah WKG, Frean M, Valera A (2022) An online offline framework for anomaly scoring and detecting new traffic in network streams. IEEE Trans Knowl Data Eng 34(11):5166–5181. https://doi.org/10.1109/TKDE.2021.3050400
8. Pelati A, Meo M, Dini P (2022) Traffic anomaly detection using deep semi-supervised learning at the mobile edge. IEEE Trans Veh Technol 71(8):8919–8932. https://doi.org/10.1109/TVT.2022.3174735
9. Ullah I, Mahmoud QH (2022) Design and development of RNN anomaly detection model for IoT networks. IEEE Access 10:62722–62750. https://doi.org/10.1109/ACCESS.2022.3176317
10. Brandão Lent DM, Novaes MP, Carvalho LF, Lloret J, Rodrigues JJPC, Proença ML (2022) A gated recurrent unit deep learning model to detect and mitigate distributed denial of service and portscan attacks. IEEE Access 10:73229–73242. https://doi.org/10.1109/ACCESS.2022.3190008
11. Zhang C, Costa-Pérez X, Patras P (2022) Adversarial attacks against deep learning-based network intrusion detection systems and defense mechanisms. IEEE/ACM Trans Netw 30(3):1294–1311. https://doi.org/10.1109/TNET.2021.3137084
12. Haydari A, Zhang M, Chuah C-N (2021) Adversarial attacks and defense in deep reinforcement learning (DRL)-based traffic signal controllers. IEEE Open J Intell Transport Syst 2:402–416. https://doi.org/10.1109/OJITS.2021.3118972

13. Xu W, Jang-Jaccard J, Singh A, Wei Y, Sabrina F (2021) Improving performance of autoencoder-based network anomaly detection on NSL-KDD dataset. IEEE Access 9:140136–140146. https://doi.org/10.1109/ACCESS.2021.3116612

14. Liu X et al (2020) NADS-RA: Network anomaly detection scheme based on feature representation and data augmentation. IEEE Access 8:214781–214800. https://doi.org/10.1109/ACCESS.2020.3040510

15. Dao T-N, Lee H (2022) JointNIDS: efficient joint traffic management for on-device network intrusion detection. IEEE Trans Veh Technol 71(12):13254–13265. https://doi.org/10.1109/TVT.2022.3198266

16. Sayed MSE, Le-Khac N-A, Azer MA, Jurcut AD (2022) A flow-based anomaly detection approach with feature selection method against ddos attacks in SDNs. IEEE Trans Cognitive CommunNetw 8(4):1862–1880. https://doi.org/10.1109/TCCN.2022.3186331

17. Wang C, Liu J (2021) An efficient anomaly detection for high-speed train braking system using broad learning system. IEEE Access 9:63825–63832. https://doi.org/10.1109/ACCESS.2021.3074929

18. Abdelmoumin G, Rawat DB, Rahman A (2022) On the performance of machine learning models for anomaly-based intelligent intrusion detection systems for the internet of things. IEEE Internet Things J 9(6):4280–4290. https://doi.org/10.1109/JIOT.2021.3103829

19. Duan G, Lv H, Wang H, Feng G (2023) Application of a dynamic line graph neural network for intrusion detection with semisupervised learning. IEEE Trans Inf Forensics Secur 18:699–714. https://doi.org/10.1109/TIFS.2022.3228493

20. Zhao J, Deng F, Li Y, Chen J (2022) Attract–repel encoder: learning anomaly representation away from landmarks. IEEE Trans Neural Netw Learn Syst 33(6):2466–2479. https://doi.org/10.1109/TNNLS.2021.3105400

21. Monshizadeh M, Khatri V, Gamdou M, Kantola R, Yan Z (2021) Improving data generalization with variational autoencoders for network traffic anomaly detection. IEEE Access 9:56893–56907. https://doi.org/10.1109/ACCESS.2021.3072126

22. Li Yet al (2022) NIN-DSC: a network traffic anomaly detection method based on deep learning. In: 7th International conference on signal and image processing (ICSIP), Suzhou, China, pp 390–394. https://doi.org/10.1109/ICSIP55141.2022.9886658

23. Sun Y, Ochiai H, Esaki H (2021) Deep learning-based anomaly detection in LAN from raw network traffic measurement. In: 2021 55th Annual conference on information sciences and systems (CISS), Baltimore, MD, USA, pp 1–5. https://doi.org/10.1109/CISS50987.2021.9400241

24. Reddy KP, Kodati S, Swetha M, Parimala M, Velliangiri S (2021) A hybrid neural network architecture for early detection of DDOS attacks using deep learning models. In: 2nd International conference on smart electronics and communication (ICOSEC), Trichy, India, pp 323–327. https://doi.org/10.1109/ICOSEC51865.2021.9591969

25. McKinney E, Mortensen D (2021) Deep anomaly detection for network traffic. In: 55th Asilomar conference on signals, systems, and computers, Pacific Grove, CA, USA, pp 1279–1283. https://doi.org/10.1109/IEEECONF53345.2021.9723308

26. Raju D, Sawai S, Gavel S, Raghuvanshi AS (2021) Development of anomaly-based intrusion detection scheme using deep learning in data network. In: 12th International conference on computing communication and networking technologies (ICCCNT), Kharagpur, India, pp 1–6. https://doi.org/10.1109/ICCCNT51525.2021.9579510

"Multi-Scale Memory Residual Network-Based Deep Learning Model for Network Traffic Anomaly Detection"
Original Submission

Evidence of Review

Reviewer/s Recommendation Term:	Major revisions necessary

Comments to Author:

1. The security analysis provided is comprehensive, but it could benefit from a clearer summary that directly states whether the protocol meets each security goal (e.g., resistance to specific attacks such as man-in-the-middle).
2. While comparisons to other protocols are mentioned, adding a detailed table summarizing the advantages over each compared protocol could visually strengthen the argument about the proposed protocol's efficiency and security.
3. More detailed pseudocode for the key algorithms (MASK and UNMASK) would enhance the reader's understanding of how these algorithms contribute to security and how they are implemented within the protocol.
4. Expanding on the performance metrics beyond storage and communication overhead to include computational overhead and power consumption would provide a more holistic view of the protocol's suitability for IoT devices.
5. The section would benefit from more details on the experimental setup, including the types of IoT devices used for testing, to better understand the context and applicability of the protocol.
6. The formal security proofs are mentioned briefly. Providing a step-by-step breakdown of these proofs can help validate the protocol's security claims more rigorously.
7. Including more diagrams or flowcharts could visually break down the protocol operation stages, particularly how the PUF-based key generation interacts with other elements of the protocol.
8. Suggesting areas for future research, such as potential improvements in the protocol or its adaptation for different types of IoT networks, would be insightful for readers looking to build on this work.
9. Discussing potential real-world applications in more depth, including specific IoT scenarios where this protocol could be particularly beneficial, would highlight its practical importance and relevance to the industry.
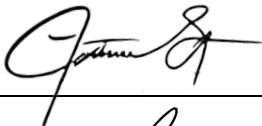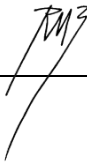
CERTIFICATION OF PERCENTAGE CONTRIBUTION

(Research Output with Multiple Authors)

Title of Research: Multi-Scale Memory Residual Network Based Deep Learning Model for Network Traffic Anomaly Detection

Type of Research Output: Book Chapter (Scopus-Indexed, Publisher: Springer Nature)

Instruction: Supply ALL the names of the authors involved in the publication of Research output and indicate the contribution of each author in percentage. Each author shall sign the conforme column if he/she agrees with the distribution. The conforme should be signed by all the authors in order to be considered. Please prepare separate Certification for each output.

	Name of Authors	Current Affiliation	% Contribution	Conforme *(Sign if you agree with the % distribution)*
1	M. Jayakrishna	Sri Sivani Engineering College	20%	
2	V. Selvakumar	Vivekanada College of Science	20%	
3	Atul Kumar	Patil B-School	20%	
4	Salunke Mangesh Dilip	GHRCEM	20%	
5	Renato R. Maaliw III	SLSU	20%	
	Should have a total of 100%		100.00%	

Prepared by:

Renato R. Maaliw III, DIT
(Name and Signature)
Faculty

Certified by:

Nicanor L. Guinto, Ph.D
(Name and Signature)
Director, Office of Research Services