2025 2nd International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE) | 979-8-3315-9848-8/25/$31.00 ©2025 IEEE | DOI: 10.1109/RMKMATE64874.2025.11042586

Zero Day Attack Prediction Using Improved Deep Neural Network

Swapna G Nair[1, a]
Research Scholar,
Department of Computer Science and Engineering,
Saveetha School of Engineering
Saveetha Institute of Medical and Technical Sciences
Saveetha University, Thandalam, Chennai, India.

[a)] swapnamgnair@gmail.com

A Jaya Mabel Rani[2, b]
Associate Professor
Department of Computer Science and Engineering
Saveetha School of Engineering
Saveetha Institute of Medical and Technical Sciences
Saveetha University, Thandalam, Chennai, India

[b)] jayamabelrania.sse@saveetha.com

Abstract—Zero-day attacks (ZDA) seriously compromise security by using unidentified weaknesses and leaving systems exposed until they are found and fixed. Usually conventional defensive systems lack real-time detection of such threats, this RESEARCH presents a fresh method based on Improved Deep Neural Network (IDNN) architecture for ZDA prediction. Combining a multi-layered deep learning framework with ideal hyperparameter tuning methods improves feature extraction and anomaly detection in the proposed model. The IDNN learns patterns suggestive of possible ZDAs using a large dataset of network traffic, therefore allowing proactive identification before major harm occurs. Using Attention Mechanisms and Residual Learning to manage the complexity of big-scale data and reduce false positives helps the model be more efficient. Extensive studies show that the suggested model provides a strong solution for real-time ZDA avoidance because it beats current techniques in accuracy, precision, and recall. This study offers a vital first step towards more strong and robust cyber security systems able to resist developing risks.

Keywords: Anomaly Detection, Deep Learning, Improved Deep Neural Network, Real-Time Detection, Zero-Day Attacks

I. INTRODUCTION

In cyber security, a basic idea is predicting zero-day (ZD) incidents. Most people and businesses lack knowledge of what their systems, networks, software, databases, or websites are vulnerable to prior to a hack [1], [2]. There have been numerous losses as well as financial ones because we cannot predict when attackscan hit [3][21]. To protect System and cyber users from zero-day attacks (ZDA) [4], [5], proactive prediction and defense systems capable of making smart judgments and predictions in real time are very vital. Two basic bases of attack prediction are statistical techniques and computational approaches. Ordinary least square regression, logistic regression, time-series techniques, auto regression, and data mining are all statistical models, while the probabilistic model and machine learning (ML)are examples of algorithmic approaches [6], [7].

When it comes to computer security, ZDAs are among the most formidable challenges. These attacks take advantage of software or hardware flaws that developers are not aware of, letting bad guys inside systems before they can fix them [8], [9]. When it comes to ZDAs, standard security measures like firewalls, intrusion detection systems, and antivirus software just do not cut it [10], [11]. Organizations confront substantial dangers, including as data breaches, monetary losses [25], and interruptions to operations, due to the rising frequency and complexity of these attacks. In order to tackle these issues, there is an increasing need for sophisticated prediction models that can identify and counteract ZDAs as they happen [12], [13]. Anomaly identification and the categorization of yet undiscovered dangers are two areas where ML and deep learning (DL) approaches have shown remarkable potential [14] [24]. Overfitting, high false-positive rates, and the large dimensionality of network traffic data are problems that plague many current models, reducing their usefulness in real-world applications [15], [16].

The main contribution of the paper is:

- Zero day attack detection using IDNN

For the duration of this paper, its structure is as follows. Several authors examine various methods for identifying ZD threats in Section 2. Section 3 displays the proposed model. In Section 4, we go over the study's results. The concluding portion of portion 5 discusses the results and ideas for future study.

A. Motivation of the paper

There is a serious vulnerability in the present state of cybersecurity that is being exposed by the growing number and complexity of ZDAs. Traditional security solutions are insufficient and reactive because these attacks take advantage of flaws that were not there before. Current approaches often fail to identify these vulnerabilities in real-time, resulting in substantial potential harm prior to the implementation of a remedy. In order to prevent damage from ZDAs, our study is driven by the need for a proactive and effective response. This study presents the Improved Deep Neural Network (IDNN) architecture, which intends to optimize hyperparameters, perform enhanced feature extraction, and identify anomalies in order to defeat the shortcomings of conventional methods. An important step in building more robust and safe cybersecurity systems, the IDNN's use of Attention Mechanisms and Residual Learning substantially improves its capacity to handle complicated data while minimizing false positives.

II. Background Study

A.O. David and O. O. Oluwasola[2] Using the dataset, the created model has a better accuracy of about 92%. By accurately predicting when attacks will occur, security experts can put safeguards in place to lessen the impact of potential threats.

Bherde, G. P., & Pund, M. A. [6]In order to fine-tune the detection system and avoid false positives during detection, a web services administrator might investigate the instructive signals.

H. Hindyet al. [7] these authors research presents new work that suggests a method for detecting ZD cyberattacks based on outliers. The most important goal was to defeat the limits of existing IDS and create an intelligent model that can detectZD cyber-attacks with a high detection accuracy. An autoencoder model for ZDA detection is proposed and tested in this work. Auto encoders' encoding-decoding capabilities served as inspiration for the concept.

M. Sarhanet al. [9] To assess how well ML-based NIDSs detect unknown attacks, or ZDAs, a new ZSL-based paradigm has been suggested. Attribute learning is the process by which the model learns, from a collection of known attacks, the characteristics that make attack traffic unique. To do this, we establish a relationship map between the characteristics of the network data and meaningful properties. To identify a ZDA, the model must establish a connection between known attack behaviors during the inference step.

P. Li et al.[12] Using meta-learning, we provide RETSINA, a new framework for detecting ZD Web attacks across many domains with insufficient training data.In order to do information sharing efficiently, we create a number of novel designs. We test RETSINA on four real-world datasets and show that it works well for a variety of diverse web domains even when training data is scarce.

R. Shad et al. [13] ZDA Detection with Unsupervised Anomaly Detection strongly argues for using unsupervised anomaly detection techniques to find ZDAs. Results of this research help to clarify the degree of effectiveness of this approach and offer the foundation for next industry growth.

S. U. Rehmanet al. [15] whether a business is vulnerable to spam or intrusion detection systems, ZDAs are a major concern. In order to get sensitive information from their rivals, malicious firms use ZD vulnerabilities that their rivals have discovered. Instead of damaging the system, attackers often aim to stay within it undetected in order to steal sensitive data[22].

V. T. Emmah et al. [16] a ZD vulnerability is a security weakness in a program or network that neither the user nor the manufacturer is aware of, and for which no fix has been made public. It is crucial to recognize the significant responsibilities that malware plays in exploiting these vulnerabilities and launching ZDAs. In this work, we have examined malware's activities from the perspective of how it attacks both software and hardware systems[23].

A. Problem definition

ZDAs, which reveal systems before patches are released, can exploit system flaws. Usually worthless against these attacks, conventional security measures are reactive and unable to quickly detect new threats. The problem is the lack of proactive, real-time detection technologies able to identify ZDAs before they do significant damage.

III. Materials and Methods

Emphasizing the Improved Deep Neural Network (IDNN) architecture, we provide the suggested strategy for ZDA prediction in this part. By combining cutting-edge methods in feature extraction and anomaly detection, the IDNN is meant to improve detection powers. The model efficiently manages and analyzes vast-scale network traffic data using a multi-layered deep learning architecture with optimal hyperparameter tuning.

A. Dataset collection

This study's dataset taken from Kaggle, a popular website for data science and ML contests. Specifically, it is available at the following URL: https://www.kaggle.com/code/mkashifn/celosia-zero-day-attack-detection-demo. This dataset is a great tool for academics and professionals working in cybersecurity as it are part of a demonstration this research aiming on ZDA detection.

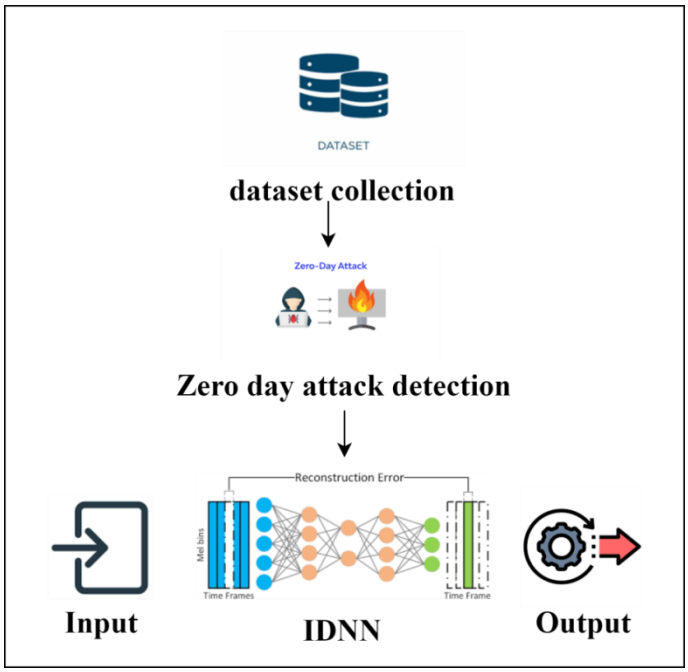

Figure 1: Proposed workflow architecture

B. Zero day attack prediction using IDNN

b. DNN

ZDA prediction using a DNN is essentially about using a deep learning model to find undiscovered vulnerabilities before they are used. Analyzing system activity and network data, the DNN can find anomalies suggesting ZDAs. Distributed neural networks (DNNs) improve real-time threat detection even in the absence of known or resolved vulnerabilities because to their multi-layered design, which allows them learn sophisticated patterns and correlations in big datasets.

Following training the support vector machine method, the learnt characteristics are grouped using deep neural networks in the next stage, known as anomaly classification.The first step is to train the grouped patches using a support vector machine; the learned features should maximize the geometric margin and minimize the classification error rate. Every description and patch is associated with a certain class, represented as {1, -1}. If a patch is relevant to the 1 class, it belongs to the characteristics connected to anomalous video anomalies; otherwise, it belongs to the -1 class. Here is the definition of the class representation:

$D = \{(x_i, y_i) | x_i \in R^p, y \in \{-1,1\}\}$ ---------- (1)

The next step, after class definition, is to identify the hyper plane that will be used for feature training.

$wt.x - b = 0$ ---------- (2)

In addition to lowering the classification error rate, the hyperplane that forms minimizes the geometric margin maximization issue.

b. Improved DNN

ZDA prediction using an Improved Deep Neural Network (IDNN) seeks to improve traditional DNNs by means of new approaches, therefore enhancing the discovery of undiscovered vulnerabilities. Among the advanced techniques the IDNN applies to improve feature extraction and anomaly detection are resilient learning and attention processing.

Training in floating-point, then directly quantizing, and finally retraining the weights are the three stages that make up the architecture of a fixed-point DNN method. Any state-of-the-art method, including unsupervised learning and dropout, can be used for the floating-point training process. Take into consideration that the optimal floating-point weights must form the basis of fixed-point optimization. The optimization of the floating-point weight can need to be repeated with different initializations, which makes this step the most time-consuming. After floating-point training, the next step is direct quantization.

For direct quantization, the uniform quantization function defines the following function, q(•).

$q(wt) = sgn(wt).\Delta.\min\left(\left[\frac{|(wt)|}{\Delta} + 0.5\right], \frac{l-1}{2}\right)$ ---------- (3)

This is where the sign function sgn(•), the quantization step size Δ, and the number of quantization levels l are defined. Since the values of the weights might be either positive or negative, it follows that M must be an odd integer. The weights are written as $-3 \bullet \Delta, -2 \bullet \downarrow, -1 \bullet \downarrow, 0, +1 \bullet \downarrow, +2 \bullet \downarrow, +3 \bullet \Delta$, which can be expressed using three bits, when M is 7.

The following is an illustration of how the quantization step size Δ is chosen in order to reduce the L2 error, E.

$E = \frac{1}{2}(q(wt_i) - wt_i)^2$ ---------- (4)

Where, wti is the i-th weight value in floating-point format. Iterations are necessary, but the procedure is not time-consuming.

Algorithm 1: IDNN

Inputs:

- Network traffic dataset with features and labels.
- Hyperparameters for the IDNN model.
- Floating-point weight initialization.

Steps:

- Initialize and configure the IDNN architecture with attention mechanisms and residual learning.
- Train the IDNN on the dataset to optimize floating-point weights.
- Apply uniform quantization to convert floating-point weights to fixed-point representations.
- Use the quantization function q $(wt) = sgn(wt).\Delta.\min\left(\left[\frac{|(wt)|}{\Delta} + 0.5\right], \frac{l-1}{2}\right)$ is the quantization step size and M is the number of quantization levels.

Outputs:

- Trained IDNN model with quantized weights.

IV. Results and Discussion

In this research, report and evaluate the ZDA prediction models, with particular attention to DNN and IDNN performance. We evaluate their performance with respect to important criteria like accuracy, precision, recall, and F-measure.

Table 1: Performance metrics comparison table

Methods	Accuracy	Precision	Recall	F-measure
CNN	94.35	94.25	94.25	95.14
RNN	95.61	95.34	95.14	95.89
DNN	96.99	97.14	97.14	98.14
IDNN	98.24	98.24	98.27	99.01

Performance examination of many models for ZDA prediction demonstrates in table 1 that the Improved Deep Neural Network (IDNN) routinely beats the other approaches. With an accuracy of 98.24%, the IDNN is the best among all models because it shows great capacity to accurately categorize attack and non-attack events. With 98.24%, it also claims the highest accuracy, therefore confirming its effectiveness in lowering false positives—that is, where the model is less likely to wrongly classify benign events as attackers. Reflecting its ability to accurately identify a higher proportion of real ZDAs, the recall of 98.27% is the best among the examined models. At last, the IDNN achieves an F-measure of 99.01%, which aggregates recall and accuracy into a single metric, therefore stressing its general resilience in juggling these two performance criteria. This complete performance advantage emphasizes how well the IDNN predicts ZDAs, exceeding CNN, RNN, and conventional DNN methods in precisely and dependably.

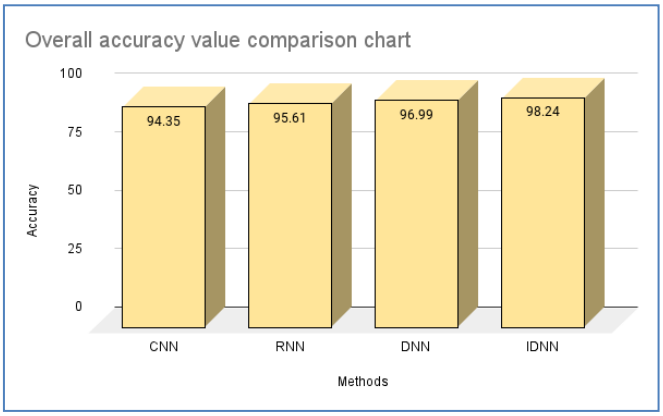

Figure 2: Overall accuracy value comparison chart

Figure 2 features a comparison table showing overall accuracy values. On the horizontal line, we can clearly see the methods; on the vertical line, the accuracy rating.

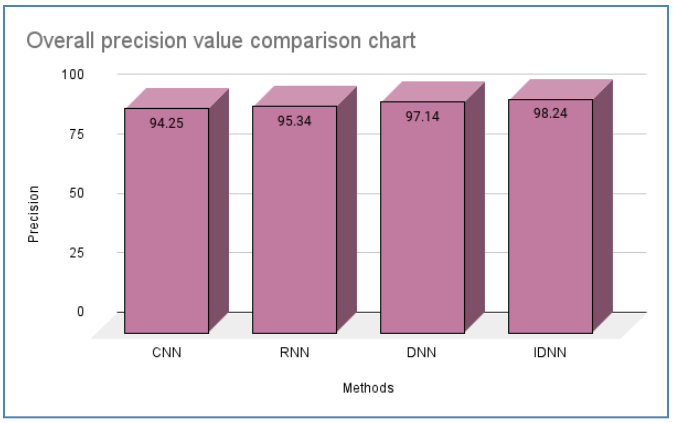

Figure 3: Overall precision value compari9son chart

Figure 3 features a comparison table showing overall accuracy values. On the horizontal lineare given methods; on the vertical lineis shown accuracy value.

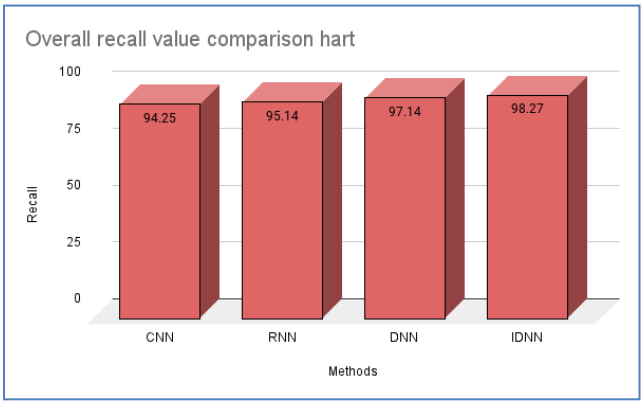

Figure 4: Overall recall value comparison chart

Figure 4 features a comparison table showing overall recall values. Procedures are represented on the horizontal linewhile recall value shown on the vertical line.

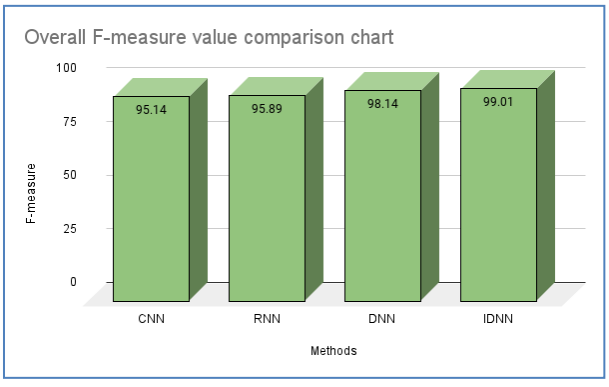

Figure 5: Overall F-measure value comparison chart

Figure five shows a comparison chart of the overall f-measure values. The procedures are indicated on the horizontal line; the f-measure value is shown on the vertical line.

V. Conclusion

Ultimately, by using cutting-edge deep learning methods including attention mechanisms and residual learning, the proposed Improved Deep Neural Network (IDNN) architecture offers a strong solution for estimating ZDAs. The model improves feature extraction and anomaly detection, therefore allowing proactive real-time possible exploit discovery. Extensive studies show that the model beats conventional approaches in accuracy, precision, and memory, hence reducing false positives. This study offers a scalable and efficient method to reduce ZD threats, therefore boosting cybersecurity defenses and acting as a necessary step toward more robust and safe systems.

VI. References

1. A. Boualouache, B. Brik, R. Rahal, Y. Ghamri-Doudane, and S. M. Senouci, "Federated Learning for Zero-Day Attack Detection in 5G and Beyond V2X Networks," *arXiv preprint arXiv:2407.03070*, 2024.
2. A. O. David and O. O. Oluwasola, "Zero day attack prediction with parameter setting using bi direction recurrent neural network in cyber security," *Int. J. Comput. Sci. Inf. Security (IJCSIS)*, vol. 18, no. 3, pp. 111–118, 2020.
3. A. Rizzardi, S. Sicari, and A. C. Porisini, "NERO: NEural algorithmic reasoning for zeRO-day attack detection in the IoT: A hybrid approach," *Comput. & Security*, vol. 142, p. 103898, 2024.
4. B. Diloglu, "Zero-Day Attack Detection with Deep Learning in Networks," Ph.D. dissertation, National College of Ireland, Dublin, 2022.
5. B. Ibrahim Hairab, H. K. Aslan, M. S. Elsayed, A. D. Jurcut, and M. A. Azer, "Anomaly detection of zero-day attacks based on CNN and regularization techniques," *Electronics*, vol. 12, no. 3, p. 573, 2023.
6. G. P. Bherde and M. A. Pund, "Technique for Detecting Zero Day Attack by using Signature based and Knowledge Based Method," 2018.
7. H. Hindy, R. Atkinson, C. Tachtatzis, J. N. Colin, E. Bayne, and X. Bellekens, "Utilising deep learning techniques for effective zero-day attack detection," *Electronics*, vol. 9, no. 10, p. 1684, 2020.
8. H. Hindy, R. Atkinson, C. Tachtatzis, J. N. Colin, E. Bayne, and X. Bellekens, "Towards an effective zero-day attack detection using outlier-based deep learning techniques," *arXiv preprint arXiv:2006.15344*, 2020.
9. M. Sarhan, S. Layeghy, M. Gallagher, and M. Portmann, "From zero-shot machine learning to zero-day attack detection," *Int. J. Inf. Security*, vol. 22, no. 4, pp. 947–959, 2023.
10. M. Sayduzzaman, J. T. Tamanna, D. Kundu, and T. Rahman, "Interoperability and Explicable AI-based Zero-Day Attacks Detection Process in Smart Community," *arXiv preprint arXiv:2408.02921*, 2024.
11. M. Tokmak, "Deep forest approach for zero-day attacks detection," in *Innovations and Technologies in Engineering*, pp. 45–56, 2022.
12. P. Li *et al.*, "Learning from Limited Heterogeneous Training Data: Meta-Learning for Unsupervised Zero-Day Web Attack Detection across Web Domains," in *Proc. 2023 ACM SIGSAC Conf. on Computer and Communications Security*, Nov. 2023, pp. 1020–1034.
13. R. Shad, A. Olukemi, and A. Egon, "Zero-Day Attack Detection with Unsupervised Anomaly Detection," 2024.
14. S. Rana, M. A. Hossan, and A. Adel, "Cloud Zero-Day Attack Detection Using Hidden Markov Model with Transductive Learning," 2021.
15. S. U. Rehman, S. Ali, G. Adeem, S. Hussain, and S. S. Raza, "Computational Intelligence Approaches for Analysis of the Detection of Zero-day Attacks," *Univ. of Wah J. Sci. Technol. (UWJST)*, vol. 6, pp. 27–36, 2022.
16. V. T. Emmah, C. Ugwu, and L. N. Onyejegbu, "An Enhanced Classification Model for Likelihood of Zero-Day Attack Detection and Estimation," *Eur. J. Electr. Eng. Comput. Sci.*, vol. 5, no. 4, pp. 69–75, 2021.
17. J. M. Rani, A. Antony *et al.*, "A big data scheme for heart disease classification in map reduce using jellyfish search flow regime optimization enabled Spinalnet," *Pacing Clin. Electrophysiol.*, vol. 47, no. 7, pp. 953–965, 2024.
18. R. Geetha, A. Saranya, K. Vijayakumar, and S. Prabha, "Classification of Retinal Fundus Images into Normal / Diabetic with Fused Deep Features," in *2024 Ninth Int. Conf. on Science Technology Engineering and Mathematics (ICONSTEM)*, Chennai, India, 2024, pp. 1–5, doi: 10.1109/ICONSTEM60960.2024.10568867.
19. V. V. B. Raju and T. P. Anithaashri, "Improvisation of data security in mobile based data by using novel blowfish algorithm over DES algorithm," in *AIP Conf. Proc.*, vol. 2729, no. 1. AIP Publishing, 2024.
20. K. Malathi, E. D. K. Ruby, and K. Gangadharan, "Revolutionizing Deep Vein Thrombosis (DVT) Management: Machine Learning Unveils Precision in Early Detection," in *2024 Ninth Int. Conf. on Science Technology Engineering and Mathematics (ICONSTEM)*, 2024, p. 15, doi: 10.1109/ICONSTEM60960.2024.10568667.
21. K.A. Mohamed Junaid, T. Sethukarasi, M. Vigilson Prem, Adi Alhudhaif, Norah Alnaim, "A novel efficient Rank-Revealing QR matrix and Schur decomposition method for big data mining and clustering (RRQR-SDM)", Information Sciences, Volume 657, 2024, 119957, ISSN 0020-0255, https://doi.org/10.1016/j.ins.2023.119957.
22. Praveen, D.S., Smart traffic management system in metropolitan cities. J Ambient Intell Human Comput 12, 7529–7541 (2021). https://doi.org/10.1007/s12652-020-02453-6
23. Bhagavathi Hariharan, "A hybrid framework for job scheduling on cloud using firefly and BAT algorithm", International Journal of Business Intelligence and Data MiningVol. 15, No. 4, pp 388-407https://doi.org/10.1504/IJBIDM.2019.102811
24. N. R. Rejin Paul, "Enhanced Trust Based Access Control for Multi-Cloud Environment", Computers, Materials & Continua DOI:10.32604/cmc.2021.018993
25. B. Hariharan, "WBAT Job Scheduler: A Multi-Objective Approach for Job Scheduling Problem on Cloud Computing", Journal of Circuits, Systems and ComputersVol. 29, No. 06, 2050089 (2020), https://doi.org/10.1142/S0218126620500899