Computer Systems Science & Engineering
DOI: 10.32604/csse.2023.030362
Article

Tech Science Press

Intelligent Cybersecurity Classification Using Chaos Game Optimization with Deep Learning Model

Eatedal Alabdulkreem[1], Saud S. Alotaibi[2], Mohammad Alamgeer[3,4], Radwa Marzouk[5], Anwer Mustafa Hilal[6,*], Abdelwahed Motwakel[6], Abu Sarwar Zamani[6] and Mohammed Rizwanullah[6]

[1]Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia
[2]Department of Information Systems, College of Computing and Information System, Umm Al-Qura University, Saudi Arabia
[3]Department of Computer Science, College of Science & Art at Mahayil, King Khalid University, Saudi Arabia
[4]Department of Computer Science and Bioinformatics, Singhania University, Pacheri Bari, Jhnujhunu, Rajasthan, India
[5]Department of Information Systems, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia
[6]Department of Computer and Self Development, Preparatory Year Deanship, Prince Sattam bin Abdulaziz University, AlKharj, Saudi Arabia
*Corresponding Author: Anwer Mustafa Hilal. Email: a.hilal@psau.edu.sa

Received: 24 March 2022; Accepted: 17 May 2022

Abstract: Cyberattack detection has become an important research domain owing to increasing number of cybercrimes in recent years. Both Machine Learning (ML) and Deep Learning (DL) classification models are useful in effective identification and classification of cyberattacks. In addition, the involvement of hyper parameters in DL models has a significantly influence upon the overall performance of the classification models. In this background, the current study develops Intelligent Cybersecurity Classification using Chaos Game Optimization with Deep Learning (ICC-CGODL) Model. The goal of the proposed ICC-CGODL model is to recognize and categorize different kinds of attacks made upon data. Besides, ICC-CGODL model primarily performs min-max normalization process to normalize the data into uniform format. In addition, Bidirectional Gated Recurrent Unit (BiGRU) model is utilized for detection and classification of cyberattacks. Moreover, CGO algorithm is also exploited to adjust the hyper parameters involved in BiGRU model which is the novelty of current work. A wide-range of simulation analysis was conducted on benchmark dataset and the results obtained confirmed the significant performance of ICC-CGODL technique than the recent approaches.

Keywords: Deep learning; chaos game optimization; cybersecurity; chaos game optimization; cyberattack

1 Introduction

Cybercrimes are increasing at an alarming rate on a daily basis which brings a disturbing directive for cybersecurity experts and specialists [1]. Hence, sophisticated board instruments that hold the capacity to recognize and forestall such occurrences in a convenient and smart manner are desperately required. The general public safety of a nation depends in this scenario. In this paper, the focus is shed upon brilliant cybersecurity frameworks or strategies to work in a smart way so as to secure the board. Normally, cybersecurity is described as an assortment of innovations and cycles that is intended to safeguard the PCs, organizations, projects, and information against malicious exercises, assaults, hurt, or unapproved access [2]. As per the existing requirements, ordinary security arrangements like antivirus, firewalls, client validation, encryption, and so forth may not be successful [3,4]. Information-driven learning strategies have developed in a quick manner in recent years to ensure cybersecurity. On a daily basis, many new malware (Malicious Software) assaults occur upon PCs and networks while most of the attacks are addressed or found at a later time while some are left out. In the past two decades, AI approaches are adjusted to the space of malware recognition/characterization which is endeavored towards intercommunication for better treatment of malware attacks as hard as zero-day attacks [5]. Lately, deep learning approaches are also involved to overcome the attacks done by malware variants [6].

Deep Learning (DL) systems turned into a functioning field to identify the intrusions that occurred in network as a part of cybersecurity [7]. While numerous studies have been conducted in this regard, there has been no studies conducted upon deep learning models, particularly on real-time datasets for intrusion location, in a controlled setting. In today's digital world, cybersecurity is a basic issue to handle [8,9]. IDS examines the network traffic or a particular PC environment to identify any malicious activity [10]. The fast development in Artificial Intelligence (AI) has brought about significant advancements in devices that include design acknowledgment and unique identification.

Ullah et al. [11] presented an integrated DL technique for detection of cyberattacks in IoT. TensorFlow DNN was presented in this study to identify the pirated software utilizing plagiarism source code. Both tokenization and weighting feature approaches were utilized to filter the noisy data. Further, the important aspects of all the tokens were focused in terms of source code plagiarism. Afterward, DL technique was utilized to detect the source code plagiarism. The authors in [12] designed an adaptive DL technique to achieve cybersecurity in which the authors enabled the recognition of attacks from social IoT. The performance of deep learning method is related to typical ML technique, and distributed attack recognition was performed against centralized recognition model.

Mihoub et al. [13] presented a novel structure integrating two elements such as DoS/DDoS detection and DoS/DDoS mitigation. The detection element offers fine-granularity recognition, as it classifies particular kind of attacks, and utilizes packet type from the attacks. In [14], advanced ML approaches were employed to detect the cyberattacks for conducting the paradigm and verification. In this study, unique test conditions were followed on several defects from GIS to demonstrate the efficiency of the presented IoT structure. The partial discharge pulse sequence features were removed from all the defects to represent the input to IoT structure.

The current study develops Intelligent Cybersecurity Classification using Chaos Game Optimization with Deep Learning (ICC-CGODL) model. The goal of the proposed ICC-CGODL model is to recognize and categorize different kinds of attacks involved in the network. Besides, ICC-CGODL model primarily performs min-max normalization process to normalize the data into a uniform format. In addition, Bidirectional Gated Recurrent Unit (BiGRU) model is also utilized for detection and classification of cyberattacks. Moreover, CGO algorithm is exploited to select the hyper parameters involved in BiGRU model. A wide-range of simulation analysis was conducted on benchmark dataset and the results were analyzed under different aspects.

2 The Proposed ICC-CGODL Model

In this study, a new ICC-CGODL technique has been developed to accomplish cybersecurity. The presented ICC-CGODL model primarily employs min-max normalization process to normalize the data into a uniform format. Followed by, BiGRU model is utilized for detection and classification of cyberattacks. Finally, CGO algorithm is exploited to choose the hyper parameters involved in BiGRU model. The workflow of the presented model is given in Fig. 1.

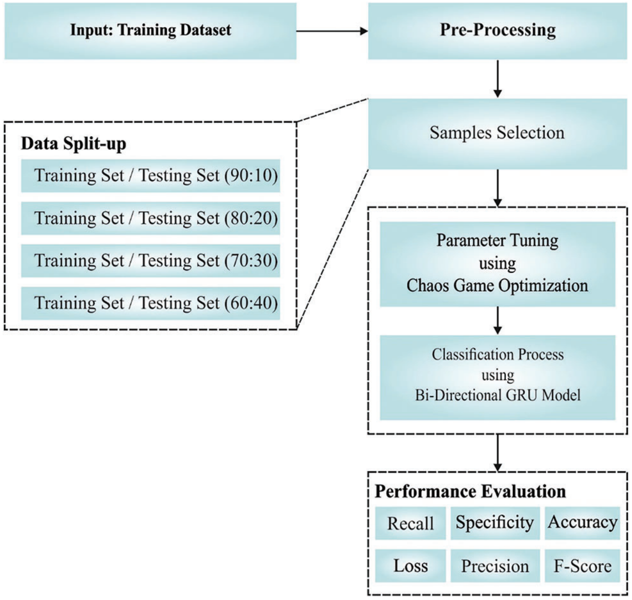
Input: Training Dataset
Pre-Processing
Samples Selection
Data Split-up
Training Set / Testing Set (90:10)
Training Set / Testing Set (80:20)
Training Set / Testing Set (70:30)
Training Set / Testing Set (60:40)
Parameter Tuning using Chaos Game Optimization
Classification Process using Bi-Directional GRU Model
Performance Evaluation
Recall
Specificity
Accuracy
Loss
Precision
F-Score

Figure 1: Working process of ICC-CGODL model

2.1 Data Pre-Processing

Primarily, the presented ICC-CGODL model employs min-max normalization process to normalize the data into a uniform format. In ML method, data normalization is usually applied to accomplish efficient presentation. The feature value varies from small to large. Thus, the normalization procedure is utilized to scale up the feature to unit variance as given below.

$$s = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2} \tag{1}$$

2.2 BiGRU Classification

After pre-processing, BiGRU model is utilized for detection and classification of cyberattacks. GRU-NN is a fundamental procedure in LSTM which is also an RNN process. GRU integrates input and forgetting gates into upgrading gates and is different from LSTM [15].

The structure of GRU model is given in Fig. 2. Let the count of hidden units be h, the small batch input is offered a time step of t being $X_t \in \mathbb{R}^{n*d}$ (the count of samples are n, the count of inputs has d), and the hidden layer (HL) at earlier time step t1 is $H_{t-1} \in \mathbb{R}^{n*h}$. The outcome HL h of single GRU at existing time step t is as follows

$$R_t = \sigma(X_t W_{xr} + H_{t-1} W_{hr} + b_r) \tag{2}$$

$$Z_t = \sigma(X_t W_{xz} + H_{t-1} W_{hz} + b_z) \tag{3}$$

$$\bar{H} = \ \tan \ h(X_t W_{xh} + (R_t E \odot H_{t-1}) W_{hh} + b_h) \tag{4}$$

$$H_t = (1 - Z_t) \odot H_{t-1} + Z_t E \odot \tilde{H}_t \tag{5}$$

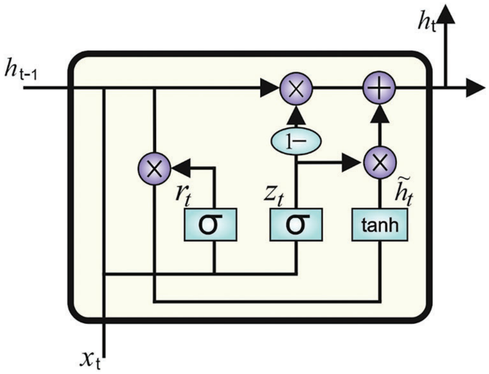

Figure 2: Structure of GRU model

In which σ signifies the sigmoid activation function, i.e., $\sigma(x) = 1/1 + e^{-x}$, b_r and b_z demonstrate the bias of reset and update gates; $\odot$ represents the matrix multiplication of two elements, W_{xr}, W_{hz}, W_{hr}, and W_{XZ} define the weights of connecting input layer and reset gate, HL and update gate, HL and reset gate, an input layer and update gate correspondingly; H_t stands for candidate HL of the existing time step t, and $Tanh$ signifies the hyperbolic tangent activation function, and the formula is as follows.

$$\tanh \ (x) = 1 - \frac{2}{1 + e^{-2x}} \tag{6}$$

When the parameter is forecasted, the value of existing time is approximately linked with the value of previous time and the value of next time. However, GRU is one-way NN infrastructure, and so Bi-GRU is employed. Bi-GRU is bi-directional NN which gathers forward- and backward-propagating GRU units. The existing HL state H_t of Bi-GRU is determined as existing input X_t, the output $\vec{H}$ of forwarded HL, and the output $\overleftarrow{\mathrm{H}_t}$ of backward HL at time step $t - 1$, then [16],

$$\overrightarrow{H}_{\mathrm{t}} = GRU(X_t, \ \overrightarrow{H}_{\mathrm{t-1}}) \tag{7}$$

$$\overleftarrow{H}_{\mathrm{t}} = GRU(X_t, \ \overleftarrow{H}_{\mathrm{t-1}}) \tag{8}$$

$$H_t = w_t \overrightarrow{H}_{\mathrm{t}} + v_t \overleftarrow{H}_{\mathrm{t}} + b_t \tag{9}$$

whereas GRU (.) function defines that GRU network is employed to conduct nonlinear change; w_t and v_t respectively are the weights of state $\overrightarrow{H}_t$ of forwarding HL and the state $\overleftarrow{H}_t$ of backward HL which is equal to Bi-GRU at time t, and b_t signifies bias.

2.3 CGO Based Hyperparameter Optimization

Lastly, CGO algorithm is exploited to choose the hyper parameters involved in BiGRU model [17,18]. CGO approach is inspired by the basic conception of chaos theory. The basic conception of chaos game and fractal elements are utilized to express a scientific model for the presented method. Due to different natural evolution procedures, a population of solution is preserved i.e., proposed by a random modification and selection procedure [17]. It is expressed in the following equation.

$$S = \begin{matrix} S_1 \\ : \\ S_n \end{matrix} = \begin{bmatrix} S_1^1 & S_1^2 & S_1^j & \cdots & S_1^d \\ S_2^1 & S_2^2 & S_2^j & & \\ & \vdots & & \ddots & \vdots \\ S_i^1 & S_i^2 & S_i^j & \cdots & S_n^d \\ S_n^1 & S_n^2 & S_n^j & & \end{bmatrix} \tag{10}$$

where $i = 1,\ 2 \ldots n. J = 1,\ 2 \ldots .\ d.$ n represents the sum of eligible seeds in Sierpinski triangle, and d signifies the dimensions. The initial location for eligible seeds is subjectively described in the searching region.

$$S_1^j\ \ (0) = \ \ S_{1,min}^j + R(S_{1,\ min}^j - S_{1,\ max}^j) \tag{11}$$

While R denotes the subjective value within [0, 1]. The initial seeds is presented herewith.

$$Seed_i^1 = S_i + x_i{*}(y_i{*}Global\ \ best - z_i{*}Mean\ \ Value) \tag{12}$$

Now $x_i,\ \ \ y_i,\ \ \ z_i$ indicate a subjective amount of 1 or 0 in probability method of rolling a dice. The presentation of the described method for the succeeding seed is given as follows.

$$Seed_i^2 = Global\ \ best + x_i{*}(y_i{*}S_i - z_i{*}Mean\ \ Value) \tag{13}$$

A presentation of 3rd and 4th seeds can be described as follows

$$Seed_i^3 = Mean\ \ Value + x_i{*}(y_i{*}S_i - z_i{*}Global\ \ best) \tag{14}$$

$$Seed_i^4 = S_i(S_i^k = S_i^k + Rand) \tag{15}$$

While k represents an arbitrary number within [0, 1]. The CGO method for x_i that controls the motion limitation of the seed [18].

$$x_i = \begin{cases} 2{*}rand \\ (\Psi{*}rand) + 1 \\ (\Omega{*}rand) + \sim \Omega \end{cases} \tag{16}$$

In which Rand indicates a subjective quantity within [0, 1]. Now, Ψ & Ω denotes an arbitrary value within [0, 1].

3 Experimental Validation

The proposed ICC-CGODL model was experimentally validated using NSL-KDD Dataset (https://www.unb.ca/cic/datasets/nsl.html) which comprises of different classes (Dos, R2l, Probe, U2r, and Normal) and a total of 41 attributes. The proposed model was simulated in MATLAB and the results are discussed herewith.

Fig. 3 shows a set of confusion matrices generated by the proposed ICC-CGODL model on distinct training/testing (TR/TS) data sizes. On TR/TS of 90:10, the proposed ICC-CGODL model recognized 4633 samples as DoS, 96 samples as R2l, 1155 samples as Probe, 0 samples as U2r, and 6619 samples as Normal. Also, on TR/TS of 80:20, the presented ICC-CGODL model classified 9064 samples as DoS, 149 samples as R2l, 2329 samples as Probe, 1 sample as U2r, and 13240 samples as Normal. Moreover, on TR/TS of 60:40, ICC-CGODL model categorized 18294 samples under DoS, 377 samples under R2l, 4548 samples under Probe, 15 samples under U2r, and 26792 samples under Normal.

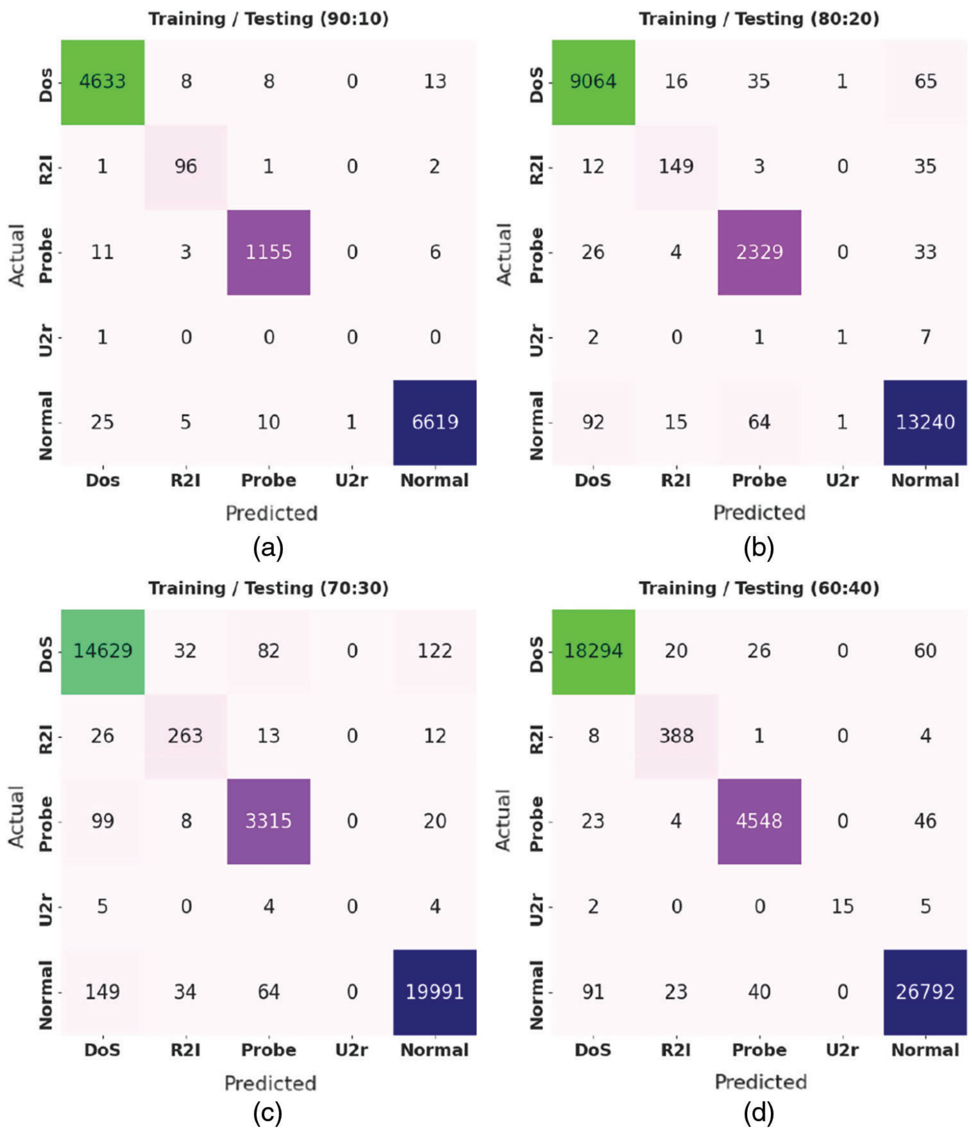

Figure 3: Confusion matrices of ICC-CGODL model

Tab. 1 and Fig. 4 report the classification results accomplished by ICC-CGODL model on TR/TS of 90:10 dataset. The results imply that ICC-CGODL model recognized DoS class with $accu_y$, $prec_n$, $reca_l$, and $spec_y$ values such as 99.47%, 99.19%, 99.38%, and 99.52% respectively. In addition, the proposed ICC-CGODL model recognized R2l class with $accu_y$, $prec_n$, $reca_l$, and $spec_y$ values such as 99.84%, 85.71%, 96%, and 99.87% respectively. Simultaneously, the ICC-CGODL model recognized Probe class with $accu_y$, $prec_n$, $reca_l$, and $spec_y$ values such as 99.69%, 98.38%, 98.30%, and 99.83% respectively. Concurrently, the ICC-CGODL model recognized Normal class with $accu_y$, $prec_n$, $reca_l$, and $spec_y$ values such as 99.51%, 99.68%, 99.38%, and 99.65% respectively.

Table 1: Classification results of ICC-CGODL model on TR/TS of 90:10

Training/Testing (90:10)				
Class labels	Accuracy	Precision	Recall	Specificity
DoS	99.47	99.19	99.38	99.52
R2I	99.84	85.71	96.00	99.87
Probe	99.69	98.38	98.30	99.83
U2r	99.98	–	–	99.99
Normal	99.51	99.68	99.38	99.65
Average	99.70	76.59	78.61	99.77

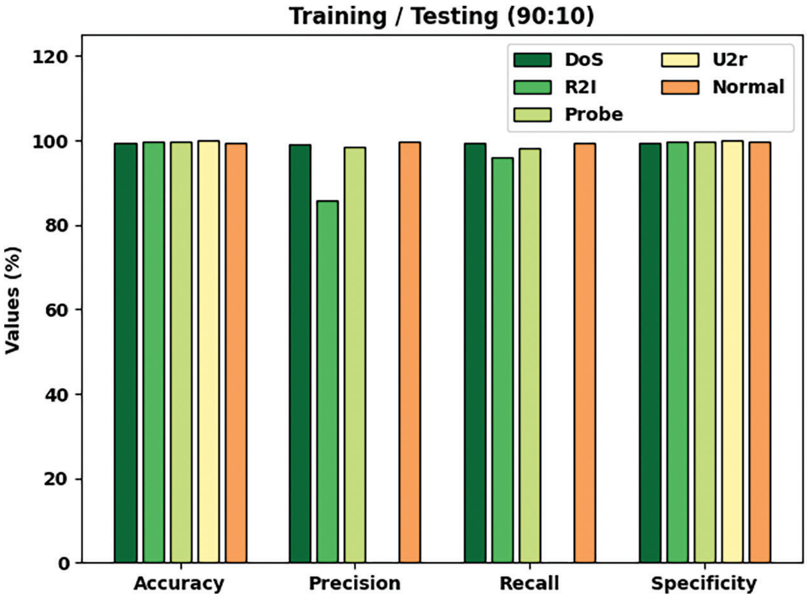

Figure 4: Cybersecurity results of ICC-CGODL model on TR/TS of 90:10

Tab. 2 and Fig. 5 examine the classification results accomplished by ICC-CGODL model on TR/TS of 80:20 dataset. The results infer that the proposed ICC-CGODL model accepted DoS class with $accu_y$, $prec_n$, $reca_l$, and $spec_y$ values such as 99.01%, 98.56%, 98.73%, and 99.18% respectively. Eventually, the ICC-CGODL model recognized R2l class with $accu_y$, $prec_n$, $reca_l$, and $spec_y$ values such as 99.66%, 80.98%, 74.87%, and 99.86% respectively. Concurrently, ICC-CGODL model recognized Probe class with $accu_y$, $prec_n$, $reca_l$, and $spec_y$ values such as 99.34%, 95.76%, 97.37%, and 99.55% respectively. Meanwhile, the ICC-CGODL model recognized Normal class with $accu_y$, $prec_n$, $reca_l$, and $spec_y$ values such as 98.76%, 98.95%, 98.72%, and 98.81% respectively.

Tab. 3 and Fig. 6 portray the classification results achieved by ICC-CGODL model on TR/TS of 70:30 dataset. The results reveal that ICC-CGODL model accepted DoS class with $accu_y$, $prec_n$, $reca_l$, and $spec_y$ values such as 98.68%, 98.13%, 98.41%, and 98.84% respectively. In line with, the ICC-CGODL model recognized R2l class with $accu_y$, $prec_n$, $reca_l$, and $spec_y$ values such as 99.68%, 78.04%, 83.76%, and 99.81% respectively. Next, ICC-CGODL model recognized Probe class with $accu_y$, $prec_n$, $reca_l$, and $spec_y$ values such as 99.25%, 95.31%, 96.31%, and 99.54% respectively. Afterward, the proposed ICC-CGODL model recognized Normal class with $accu_y$, $prec_n$, $reca_l$, and $spec_y$ values such as 98.96%, 99.22%, 98.78%, and 99.15% respectively.

Table 2: Classification results of ICC-CGODL model on TR/TS of 80:20

Training/Testing (80:20)				
Class labels	Accuracy	Precision	Recall	Specificity
DoS	99.01	98.56	98.73	99.18
R2I	99.66	80.98	74.87	99.86
Probe	99.34	95.76	97.37	99.55
U2r	99.95	33.33	9.09	99.99
Normal	98.76	98.95	98.72	98.81
Average	99.35	81.52	75.75	99.48

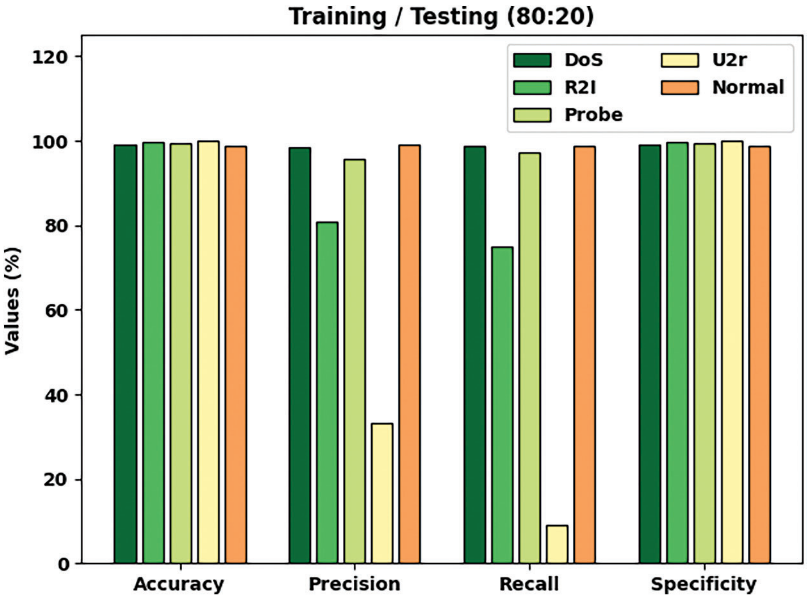

Figure 5: Cybersecurity results of ICC-CGODL model on TR/TS of 80:20

Table 3: Classification results of ICC-CGODL model on TR/TS of 70:30

Training/Testing (70:30)				
Class labels	Accuracy	Precision	Recall	Specificity
DoS	98.68	98.13	98.41	98.84
R2I	99.68	78.04	83.76	99.81
Probe	99.25	95.31	96.31	99.54
U2r	99.97	–	–	100.00
Normal	98.96	99.22	98.78	99.15
Average	99.31	74.14	75.45	99.47

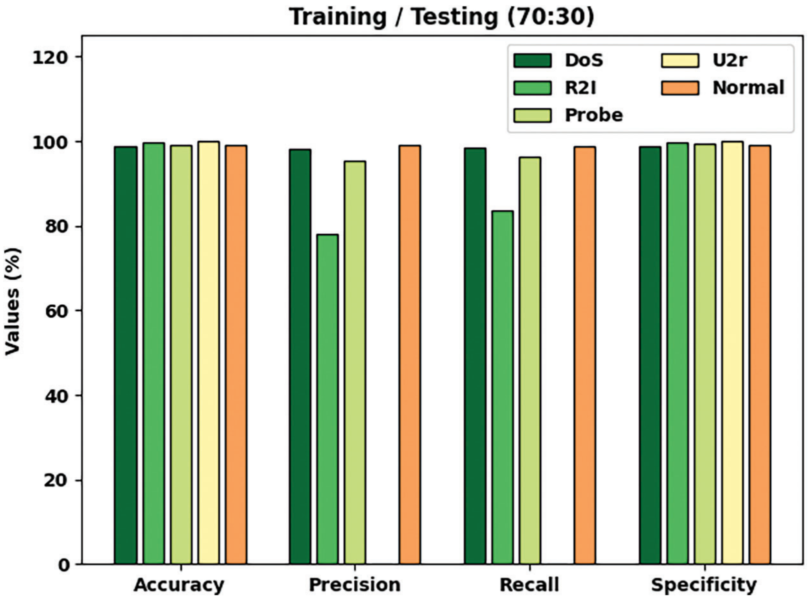

Figure 6: Cybersecurity results of ICC-CGODL model on TR/TS of 70:30

Tab. 4 and Fig. 7 portray a detailed overview on the classification results of ICC-CGODL model on TR/TS of 80:20 dataset. The results infer that the proposed ICC-CGODL model established DoS class with $accu_y$, $prec_n$, $reca_l$, and $spec_y$ values such as 99.54%, 99.33%, 99.42%, and 99.61% respectively. Eventually, ICC-CGODL model recognized R2l class with $accu_y$, $prec_n$, $reca_l$, and $spec_y$ such as 99.88%, 89.20%, 96.76%, and 99.91% respectively. Concurrently, the ICC-CGODL model recognized Probe class with $accu_y$, $prec_n$, $reca_l$, and $spec_y$ values namely, 99.72%, 98.55%, 98.42%, and 99.85%.

Table 4: Classification results of ICC-CGODL model on TR/TS of 60:10

Training/Testing (60:40)				
Class labels	Accuracy	Precision	Recall	Specificity
DoS	99.54	99.33	99.42	99.61
R2I	99.88	89.20	96.76	99.91
Probe	99.72	98.55	98.42	99.85
U2r	99.99	100.00	68.18	100.00
Normal	99.47	99.57	99.43	99.51
Average	99.72	97.33	92.44	99.78

Fig. 8 demonstrates the training/validation accuracy values achieved by ICC-CGODL approach on test dataset. The figure implies that ICC-CGODL approach produced the maximum training/validation accuracies on test data with an increase in epoch count.

Fig. 9 establishes the training/validation losses reported by ICC-CGODL approach on test dataset. The outcome infers that the proposed ICC-CGODL approach reached the least training/validation losses on test data with an increase in epoch count.

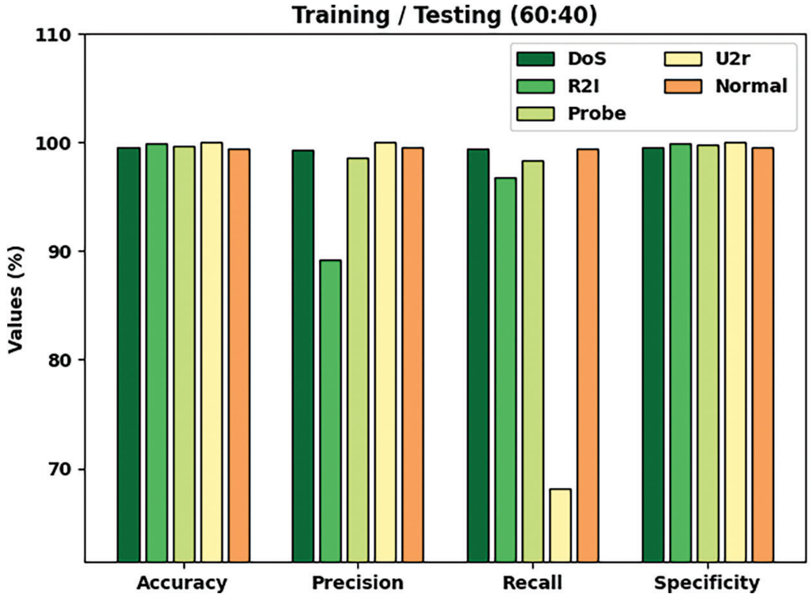

Figure 7: Cybersecurity results of ICC-CGODL model on TR/TS of 60:40

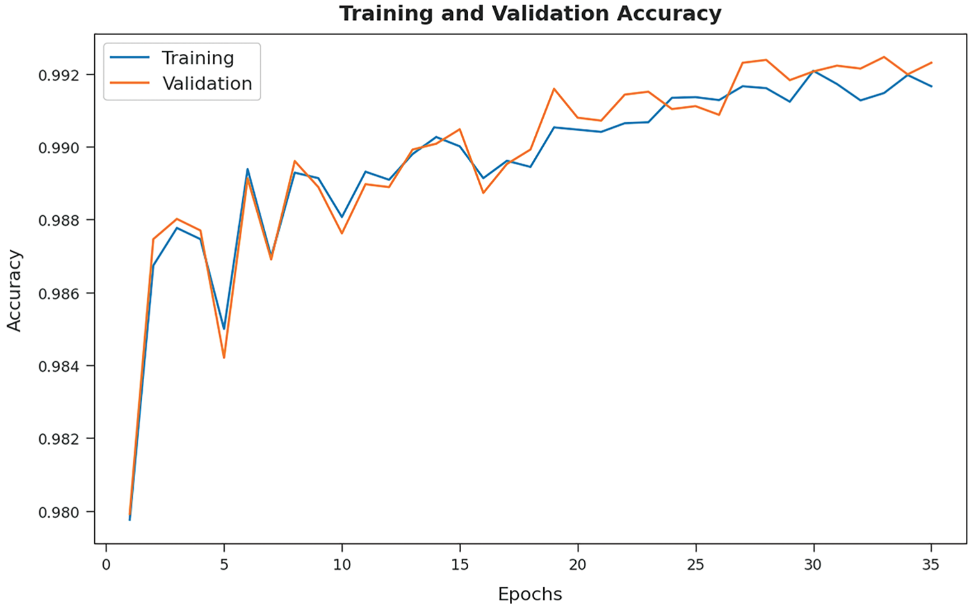

Figure 8: Training/validation accuracies of ICC-CGODL model

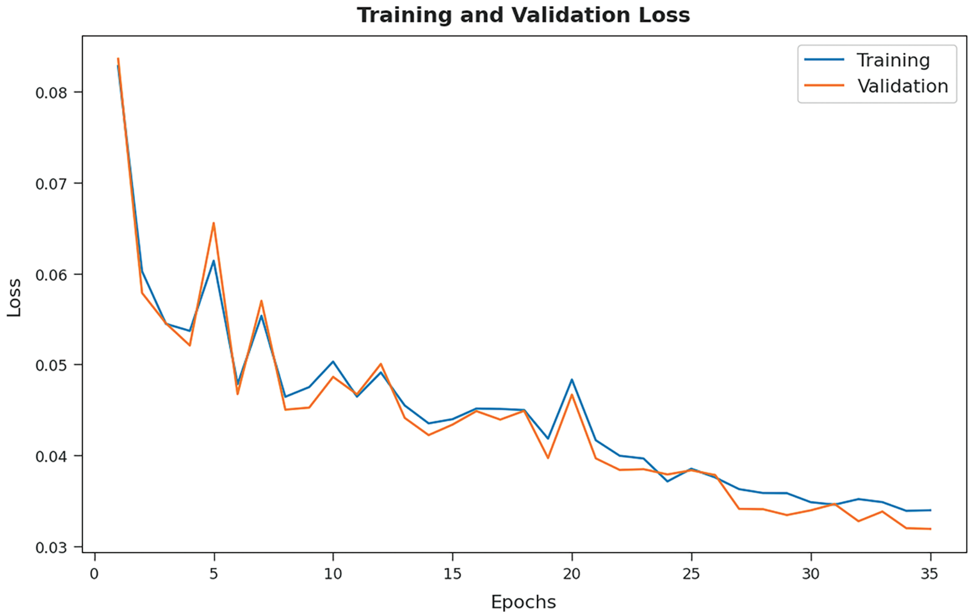

Figure 9: Training/validation losses of ICC-CGODL model

Finally, a brief comparative study was conducted between ICC-CGODL model against recent models and the results are shown in Tab. 5 and Fig. 10. The experimental results indicate that the proposed ICC-CGODL model accomplished the maximum $accu_y$ of 99.72% whereas RF, DT, SVM, NB, and KNN models attained less accuracy values such as 99.47%, 99.80%, 89.18%, 92.12%, and 98.57%. These values confirm that the proposed ICC-CGODL model outperformed existing methods.

Table 5: Comparative classification results of ICC-CGODL model

Methods	Accuracy
ICC-CGODL	99.72
Random forest model	99.47
Decision tree model	99.80
SVM model	89.18
Naïve bayes model	92.12
KNN model	98.57

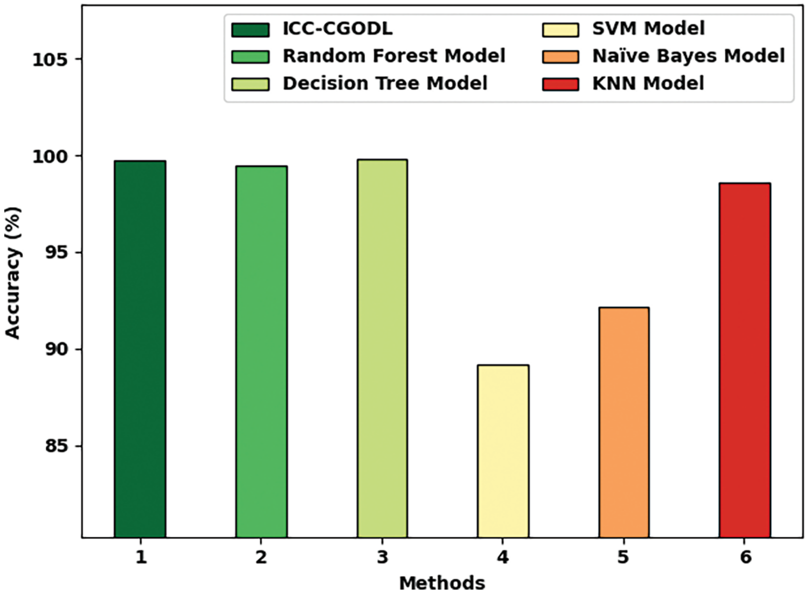

Figure 10: Comparative classification results of ICC-CGODL model

4 Conclusion

In this study, a new ICC-CGODL technique has been developed to accomplish cybersecurity. The presented ICC-CGODL model primarily employs min-max normalization process to normalize the data into a uniform format. Followed by, BiGRU model is utilized for cyberattack detection and classification. Finally, CGO algorithm is exploited to choose the hyper parameters involved in BiGRU model. A wide-range of simulation analysis was conducted on benchmark dataset and the results confirmed the significant performance of ICC-CGODL technique compared to recent approaches. Therefore, the presented ICC-CGODL technique can be employed for cyberattack detection and classification. In future, feature selection approaches can be included to decrease the computational complexity.

Funding Statement: The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work under Grant Number (RGP 2/180/43). Princess Nourah bint Abdulrahman University Researchers Supporting Project Number (PNURSP2022R161), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. The authors would like to thank the Deanship of Scientific Research at Umm Al-Qura University for supporting this work by Grant Code: (22UQU4210118DSR07).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] P. Dixit and S. Silakari, “Deep learning algorithms for cybersecurity applications: A technological and status review,” *Computer Science Review*, vol. 39, pp. 100317, 2021.

[2] D. Berman, A. Buczak, J. Chavis and C. Corbett, “A survey of deep learning methods for cyber security,” *Information*, vol. 10, no. 4, pp. 122, 2019.

[3] I. H. Sarker, “Deep cybersecurity: A comprehensive overview from neural network and deep learning perspective,” *SN Computer Science*, vol. 2, no. 3, pp. 154, 2021.

[4] Y. Xin, L. Kong, Y. Chen, Y. Li, H. Zhu *et al.,* “Machine learning and deep learning methods for cybersecurity,” *IEEE Access*, vol. 6, pp. 35365–35381, 2018.

[5] S. H. Haji and S. Y. Ameen, "Attack and anomaly detection in iot networks using machine learning techniques: A review," *Asian Journal of Research in Computer Science*, pp. 30–46, 2021.

[6] O. B. Fredj, A. Mihoub, M. Krichen, O. Cheikhrouhou and A. Derhab, "CyberSecurity attack prediction: A deep learning approach," in *13th Int. Conf. on Security of Information and Networks*, Merkez Turkey, pp. 1–6, 2020.

[7] R. Geetha and T. Thilagam, "A review on the effectiveness of machine learning and deep learning algorithms for cyber security," *Archives of Computational Methods in Engineering*, vol. 28, no. 4, pp. 2861–2879, 2021.

[8] M. A. Ferrag, L. Maglaras, S. Moschoyiannis and H. Janicke, "Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study," *Journal of Information Security and Applications*, vol. 50, pp. 102419, 2020.

[9] E. Rodriguez, B. Otero, N. Gutierrez and R. Canal, "A survey of deep learning techniques for cybersecurity in mobile networks," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 3, pp. 1920–1955, 2021.

[10] D. Chen, P. Wawrzynski and Z. Lv, "Cyber security in smart cities: A review of deep learning-based applications and case studies," *Sustainable Cities and Society*, vol. 66, pp. 102655, 2021. https://doi.org/10.1016/j.scs.2020.102655.

[11] F. Ullah, H. Naeem, S. Jabbar, S. Khalid, M. A. Latif *et al.,* "Cyber security threats detection in internet of things using deep learning approach," *IEEE Access*, vol. 7, pp. 124379–124389, 2019.

[12] A. A. Diro and N. Chilamkurti, "Distributed attack detection scheme using deep learning approach for internet of things," *Future Generation Computer Systems*, vol. 82, pp. 761–768, 2018.

[13] A. Mihoub, O. B. Fredj, O. Cheikhrouhou, A. Derhab and M. Krichen, "Denial of service attack detection and mitigation for internet of things using looking-back-enabled machine learning techniques," *Computers & Electrical Engineering*, vol. 98, pp. 107716, 2022.

[14] M. Elsisi, M. Q. Tran, K. Mahmoud, D. E. A. Mansour, M. Lehtonen *et al.,* "Towards secured online monitoring for digitalized gis against cyber-attacks based on iot and machine learning," *IEEE Access*, vol. 9, pp. 78415–78427, 2021.

[15] J. Liu, Y. Yang, S. Lv, J. Wang and H. Chen, "Attention-based BiGRU-CNN for Chinese question classification," *Journal of Ambient Intelligence and Humanized Computing*, 2019. https://doi.org/10.1007/s12652-019-01344-9.

[16] L. Zhou and X. Bian, "Improved text sentiment classification method based on BiGRU-attention," *Journal of Physics: Conference Series*, vol. 1345, no. 3, pp. 032097, 2019.

[17] S. Talatahari and M. Azizi, "Optimization of constrained mathematical and engineering design problems using chaos game optimization," *Computers & Industrial Engineering*, vol. 145, pp. 106560, 2020.

[18] S. Talatahari and M. Azizi, "Chaos game optimization: A novel metaheuristic algorithm," *Artificial Intelligence Review*, vol. 54, no. 2, pp. 917–1004, 2021.