DDoS Attack Detection using Long-short Term Memory with Bacterial Colony Optimization on IoT Environment

Latifa Alamer[1*], Ebtesam Shadadi[2]

[1*]Lecturer, Department of Information Technology and Security, Jazan University, Jazan, Kingdom of Saudi Arabia. laalamer@jazanu.edu.sa, ORCID: https://orcid.org/0000-0002-6538-8072

[2] Lecturer, Department of Computer Science, College of Computer Science & Information Technology, Jazan University, Jazan, Kingdom of Saudi Arabia. ashedadi@jazanu.edu.sa, ORCID: https://orcid.org/ 0000-0001-9684-3261

Received: October 10, 2022; Accepted: December 06, 2022; Published: February 28, 2023

Abstract

The IoT (Internet of Things) connects everyday things to devices that communicate with one another or with other systems across a network. DDoS (Distributed Denial of Service) attack is a significant security issue in the IoT environment. Detecting DDoS attacks is a difficult and important challenge for improving the performance of IoT technologies. Recently, a well-known recurrent neural network (RNN) called the long-short term memory (LSTM) has been used to identify DDoS attacks. However, as the LSTM model's parameters are frequently established by experience, subjectivity is significant and will have an impact on the model's capacity. In this research work, the parameters of LSTM are optimized by BCO (bacterial colony optimization) for making a more efficient DDoS detection method called BCO-LSTM. The performance of BCO-LSTM is compared with conventional LSTM and some enhanced LSTM. The investigational effects indicate that the proposed BCO-LSTM outperforms compared detectors at accurately capturing the dynamic behaviors of unknown network traffic.

Keywords: DDoS Attack, Detection, Long-short Term Memory, Bacterial Colony Optimization, Internet of Things.

1 Introduction

More changes in the way technology spread in society and the economy are projected as a result of the IoT. The IoT connects everyday things to devices that communicate with one another or with other systems across a network. IoT makes it possible to exploit hyper-connected environments, such as those found in homes, factories, and hospitals, to supply new sorts of services (H. Mrabet, 2020). However, numerous hurdles must be overcome to reap the full benefits of this technology, notably in terms of security and privacy (F. Ullah, 2019). The DDoS attack is a significant security issue to the Internet in the current era. The attack is reported to be targeted specifically at the IoT ecosystem, as devices are said to have less memory, computational capacity, and security measures in place to prevent DDoS attacks. DDoS attacks are severe attacks that create confusion on IoT-connected devices and

Journal of Internet Services and Information Security (JISIS), volume: 13, number: 1 (February), pp. 44-53.
DOI: *10.58346/JISIS.2023.I1.005*

*Corresponding author: Lecturer, Department of Information Technology and Security, Jazan University, Jazan, Kingdom of Saudi Arabia.

reduce network performance. It cuts down on network and computational resources like CPU, memory, and network bandwidth (A. Wani, 2020).

Deep learning (DL) methodologies have been used to identify DDoS attacks more frequently. They've gotten a lot of attention because of their high detectability (M.I. Ahamed, 2021). LSTM is a DL method that has the capability of learning longer temporal sequences. User experience is typically used to determine the LSTM network's parameters. The behavior of several components of the algorithm is highly influenced by a set of parameters used in LSTM. It should be emphasized that there isn't a universally ideal setup for any problem domain. Therefore, it is crucial to optimize LSTM parameters to get good performance for each problem domain, such as DDoS attack detection. The parameters include regularization parameters as well as optimization parameters like batch size and the number of hidden neurons, weight, and bias optimization.

In this research work, the LSTM's parameters are optimized using the BCO, which also reduces the subjective impact of manually selected parameters. The research work aims to enhance the detection rate by obtaining the optimal parameter of LSTM.

- The BCO is used to achieve the finest parameters of LSTM to build a more efficient detection scheme
- The proposed BCO-LSTM is used to study the network traffic behavior for detecting malicious data packets.
- Three performance methods are measured to analyze the performance of the proposed BCO-LSTM.
- The other sections of this study paper are systematized in the following manner: section 2 covers similar works, section 3 discusses LSTM, section 4 talks BCO, section 5 discusses BCO-LSTM, and sections 6 and 7 respectively discuss results analysis and conclusions.

2 Literature Works

DL is a broad category of machine learning (ML) and has been used to solve difficult problems in a variety of fields. This section discusses several noteworthy recent works on this topic. Several studies use DL methods to determine non-linear demonstrations from input data for classifying and distinguishing malicious traffic from genuine traffic. Z. Li et al (2019) developed a new detection technique for detecting DDoS attacks that combines attack characteristics that makes use of Tsallis entropy and ELM (extreme learning machine) (Irshad Ahamed, M., 2019). Z. Hu et al., (2020) new detection scheme using ADASYN (Adaptive synthetic sampling) and an advanced CNN (convolutional neural network). To begin, the balance of the sample distribution using the ADASYN approach effectively prevents from being sensitive to huge examples while ignoring small ones. The enhanced CNN is use the split convolution element (SPCCNN), which can boost feature variety while removing the effect of exchange information redundancy on classical training. Finally, the ASCNN is tested using the standard NSL-KDD dataset.

G. D. L. T. Parra et al. (2020) developed a cloud-based temporal LSTM for identifying attacks and ingesting CNN embedding to detect distributed phishing attacks across various IoT devices hosted on the backend. A. Churcher et al (2021) develop a comparison of ML for both binary and multi-class classification on Bot-IoT. However, the method for detecting DDoS attacks described above has the following weaknesses: these indices are influenced by the number of datasets, the longest training time, and the lowest detection accuracy.

LSTM (Long Short-Term Memory)

LSTMs differ from standard feed-forward in that they feature feedback connections (S. Hochreiter, 1997). This is an overview of a gating unit system that maintains previous information on the internal memory unit cell state unit. LSTM relies on various "gates" to allow the network to learn continuously, such as when to forget past knowledge or update cell state with recent input. Three "gates" regulate the internal memory unit, which records all past data up to the present time. The input gate regulates the amount of new data in the internal memory unit, which is indicated as:

$$g_i^t = \sigma(\sum_j U_{i,j}^t x_j^t + \sum_j W_{i,j}^g h_j^{t-1} + b_i^g) \quad (1)$$

σ is the activation function (sigmoid). x^t is the input vectors. h^t is the present hidden layer vector, which contains all LSTM cells' output. b^g, U^g, and w^g are the input gate's bias, weights of input, and recurrent, correspondingly.

The forget gate unit regulates how much information from previous moments should be saved in the internal memory unit, which is indicated as:

$$f_i^t = \sigma(\sum_j U_{i,j}^f x_j^t + \sum_j W_{i,j}^f h_j^{t-1} + b_i^f) \quad (2)$$

The sigmoid function, having a value range 0 and 1, is called σ. b^f, U^f, and w^f are the forget gate's bias, weights of input, and recurrent, correspondingly. The forget gate is opened when $f_i^t = 1$, and the previous cell state is fed into the cell. When $f_i^t = 0$ the forget gate is closed, the prior time's cell state is discarded.

The LSTM cell's internal state unit S_i^t, which contains a conditional self-loop weight f_i^t, is updated, as follows:

$$S_i^t = f_i^t S_i^{t-1} + g_i^t \sigma(\sum_j U_{i,j} x_j^t + \sum_j W_{i,j} h_j^{t-1} + b_i) \quad (3)$$

Where bias, weights of input, and recurrent are denoted by b, U, and w correspondingly. The right side of the above equation is the cell state information measured by the forget gate at the last instant, while the second fragment is the input measured by the input gate.

The output gate regulates the internal memory unit, yielding and producing required data, which can be provided by:

$$h_j^t = \tanh(S_i^t) O_i^t \quad (4)$$

$$O_i^t = \sigma(\sum_j U_{i,j}^0 x_j^t + \sum_j W_{i,j}^0 h_j^{t-1} + b_i^0) \quad (5)$$

b^0, U^0, and w^0 are the output gate's bias, input, and weights, respectively. In the RNN, the output of the LSTM unit operates as a hidden layer.

Bacterial Colony Optimization

Niu et al. (2012) developed BCO, a population-based algorithm that mimics the behavior of artificial bacteria. The advantage of the BCO method is that it allows individuals to exchange information rather

than swimming at random. There are five phases of the BCO: chemotaxis, communication, elimination, reproduction, and migration. Tumbling and swimming are the two types of chemotaxis in bacteria. A stochastic direction contributes to the actual swimming method when tumbling. As a result, the turbulence director and the best searching director together influence the search direction in tumbling, updating the locations of each bacterium, however no turbulent director allowing in the swimming method affects the bacteria swimming near optimal, as follows:

$$\begin{aligned} Position_i(t) = Position_i(t-1) + C(i) \\ *[\, f_i\, (G_{best} - Position_i(t-1)) + (1-f_i) \\ *\, (P_{best_i} - Position_i(t-1)) + turb_i\,] \end{aligned} \quad (6)$$

Algorithm 1: Proposed BCO-LSTM

Step 1: Initialize the required parameters for BCO and LSTM
Step 2: Read the data and normalize the data using a min-max method
Step 3: The training dataset and the test dataset are then created from the normalized data.
Step 4: To train the LSTM, use the training dataset.
Step 4.1: The BCO optimizes the precise parameters in the LSTM.
Step 4.2: Train the LSTM network. The loss function chosen is the mean square error (MSE). Through training iterations, update the loss function.
Step 4.3: Determine the individual best fitness value and the global optimal fitness value for each bacterial species.
Step 4.4: the best LSTM parameters are recorded when termination conditions of BCO are met. If not, go to step 4.1 to continue the iteration.
Step 5: Utilizing the test dataset, the trained model is evaluated.
Step 6: The suggested method can determine the best LSTM model parameters by the aforementioned procedure.
Step 7: For detecting DDoS attacks, the best LSTM model is utilized, and four performance metrics are examined.

$$\begin{aligned} Position_i(t) = Position_i(t-1) + C(i) \\ *[\, f_i\, (G_{best} - Position_i(t-1)) + (1-f_i) \\ *\, (P_{best_i} - Position_i(t-1))\,] \end{aligned} \quad (7)$$

$$C(i) = C_{\min} + \left(\frac{Iter_{\max} - Iter_j}{Iter_{\max}} \right) (C_{\max} - C_{\min}) \quad (8)$$

Where, $turb_i$ is signify the turbulent direction and chemotaxis size denoted by $C(i)$. $f_i \in (0,1)$. P_{best} and G_{best} are the global best and personal best correspondingly. $Iter_{\max}$, $Iter_j$ are the maximum and current iteration respectively.

It is during the elimination and reproduction phase of the process that bacteria with a high energy will reproduce to produce new individuals, while bacteria with a low energy will be exchanged for healthier bacteria. The bacterium does excellently when it comes to hunting for high-energy nutrients. When a specific probability is met, the bacteria migrate to look for new nutrients in the migration phase.

Proposed Bco-Lstm

Many research works have shown that an LSTM network may extract useful information from low-volume data and also be applied to detect DDoS attacks (X. Liang, 2019). Furthermore, LSTM is a

method that is particularly well suited to the processing of time series. Nonetheless, the LSTM parameters are usually defined by the user's experience. The model's performance will be influenced by the parameters chosen. Numerous meta-heuristics optimization methods for parameter optimization, for example, GA (genetic algorithm) (H. Chung, 2018), PSO (particle swarm optimization) (P. Wang, 2020), (IPSO) improved PSO (Y. Ji, 2021), (B. Shao, 2019), and so on, have been presented to address the aforesaid problem.

However, the GA, PSO, and their improved versions have many shortcomings such as slow convergence rate and premature convergences. The BCO is applied to various real-world applications which produced high accuracy compared with conventional algorithms (S.S. Babu, 2022), (J. Revathi, 2019). Hence, the BCO is used to find the best parameters for LSTM which can help to enhance the performance of LSTM such as detection rate and fast convergence with high accuracy. Algorithm 1 shows the steps of the proposed BCO-LSTM and Figure 1 depicts the suggested BCO-LSTM method's flowchart.

3 Experimental Results

The power of the BCO-LSTM is compared with IPSO-LSTM (Y. Ji, 2021), PSO-LSTM (P. Wang, 2020) and GA-LSTM (Y. Huang, 2021), and LSTM (X. Liang, 2019). The models are used in a computer with an Intel Core i5-10510U CPU running at 1.8 GHz with 8GB of RAM on Windows 10 and MATLAB 2015R serves as the development environment.

1) Datasets

Two datasets are considered for analyzing the performance of compared detection methods. The details are given as follows: An authentic network background was generated at UNSW to develop the BoT-IoT. In addition to botnet traffic, there was normal traffic as well. Separating the data files into attack groups and subgroups helped in the tagging technique. (N. Koroniotis, 2019). The pcap files are 69.3 GB and 72,000,000 data samples. The mined flow traffic is 16.7 GB (CSV). DoS, DDoS, OS and Service Scan, Data Exfiltration, and Key-logging attacks are involved in the dataset, with DoS and DDoS attacks being further categorized based on the protocol utilized.

Table 1: Confusions Matrix

	Predicted normal	**Predicted attack**
Actual normal	TP (true positive)	FN (false negative)
Actual attack	FP (False positive)	TN (true negative)

The CICIDS2017 dataset (I. Sharafaldin, 2019) contains the most recent and benign prevalent attacks (PCAPs) and the outcomes of a network traffic analysis performed using CIC Flow Meter. The data collection period lasted for 5 days, from Monday, July 3, 2017, from 9 AM through Friday, July 7, 2017, at 5 PM. Monday is an ordinary day with only light traffic. Tuesday, Wednesday, Thursday, and Friday saw their morning and afternoon executions.

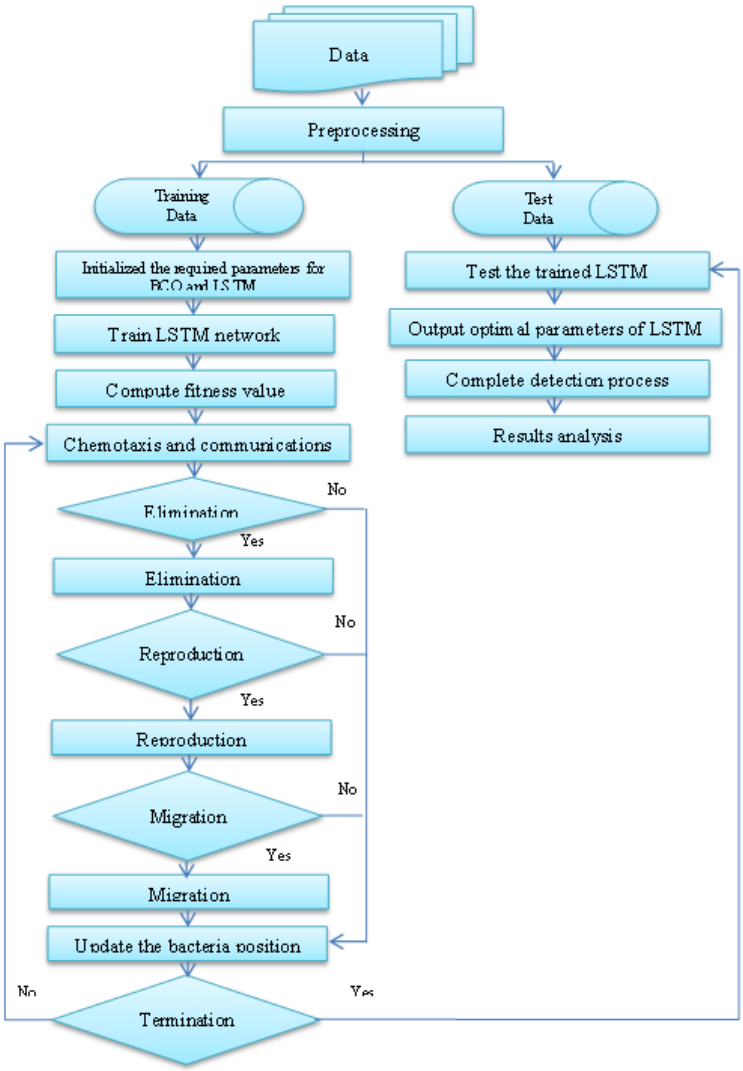

Figure 1: Flowchart for Proposed BCO-LSTM

2) Data Pre-processing

In data analytics, data preparation is a vital stage because high-quality data enables more accurate detection models. The min-max (K. Kalaiselvi, 2018), (K. Velusamy, 2017), (Irshad Ahamed, 2022) is used to normalize numbers in the range of [0, 1] which is determined as follows,

$$x'(i) = \frac{x(i) - x_{\min}}{x_{\max} - x_{\min}} \qquad (9)$$

where , $x(i)$ is i^{th} data samples, $x'(i)$ for the normalized input data derived from $x(i)$ and $x_{\max}$ and $x_{\min}$ for the biggest and smallest values of the raw data, respectively. The collected datasets are split into two sets such as training and test data. Out of the entire dataset, 70% of the samples are used as training and the remaining 30% are used as test data.

Table 2: Performance Analysis Results of BoT-IoT

Detection methods	Accuracy (%)	Specificity (%)	Sensitivity (%)
BCO-LSTM	**98.75**	**97.28**	**98.32**
IPSO-LSTM	96.82	96.82	95.28
PSO-LSTM	96.02	95.92	94.74
GA-LSTM	94.91	93.79	91.42
LSTM	90.37	91.93	90.02

Table 3: Performance Analysis Results of CICIDS2017

Detection methods	Accuracy (%)	Specificity (%)	Sensitivity (%)
BCO-LSTM	**95.38**	**97.28**	**99.76**
IPSO-LSTM	92.82	96.82	97.84
PSO-LSTM	86.25	93.72	96.73
GA-LSTM	79.18	87.67	94.84
LSTM	75.28	85.56	92.67

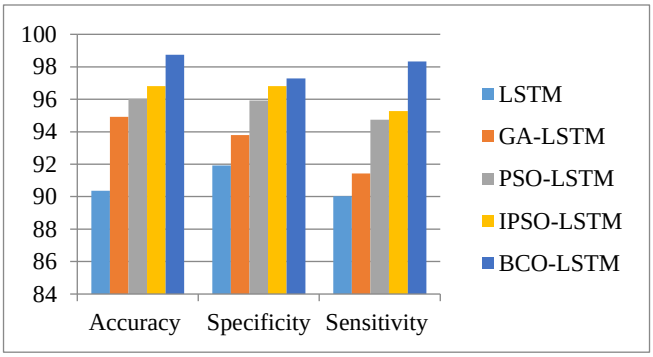

Figure 2: Performance Analysis of Detection Algorithms

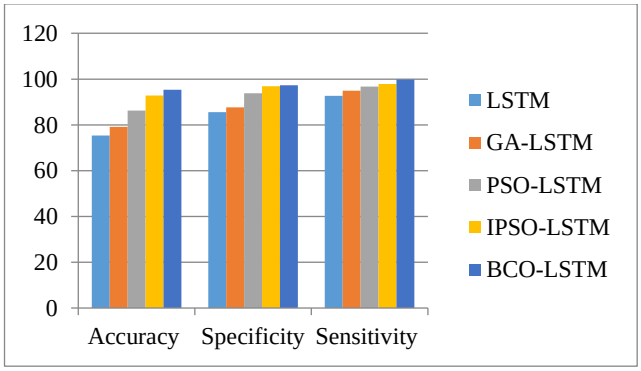

Figure 3: Performance Analysis of Detection Algorithms

3) Performance Indicators

To evaluate the prediction algorithm's performance, a variety of performance measures were employed

A. Accuracy

In order to assess the accuracy of the system, we can consider the following:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (10)$$

B. Specificity

The specificity of a prediction is determined by the percentage of actually negative cases that are categorized as negative as follows:

$$Specificity = \frac{TN}{TN + FP} \times 100\% \quad (11)$$

C. Sensitivity

In the case of true positive rates, they are calculated as follows.

$$Sensitivity = \frac{TP}{TP + FN} \times 100\% \quad (12)$$

D. Results and Discussions

Table 2 shows the experimental results based on statistical parameters for all compared algorithms in terms of the BoT-IoT. The graphical representation of performance comparisons is shown in Figure 2 For example, in terms of BoT-IoT, the highest accuracy obtained by the proposed BCO-LSTM was 98.75 %, IPSO-LSTM was 96.82 %, PSO-LSTM was 96.02 %, GA-LSTM was 94.91 %, and LSTM was 90.37 %. Similarly, according to specificity and sensitivity, the proposed BCO-LSTM produced higher performance.

Table 3 shows the experimental results based on statistical parameters for all compared algorithms in terms of the CICIDS2017. The graphical representation of performance comparisons is shown in Figure 3 For example, in terms of BoT-IoT, the highest accuracy obtained by the proposed BCO-LSTM was 95.38 %, IPSO-LSTM was 92.82 %, PSO-LSTM was 86.25 %, GA-LSTM was 87.67 %, and LSTM was 85.56 %. Similarly, according to specificity and sensitivity, the proposed BCO-LSTM produced higher performance. From overall experimental analysis, the proposed BCO-LSTM produces high accuracy and detection rate with a fast convergence rate.

4 Conclusions

A new DDoS detection scheme in the IoT environment has been released based on LSTM and is boosted by BCO. This solution addresses a major flaw in traditional parameter selection methods. The scheme's capacity to learn complicated representations automatically, allows it to distinguish between genuine and malicious communications. The results of the experiments presented that the established system had a higher detection rate and was quick to learn. However, the BCO algorithm's inner iterations and capacity for random searches cause its poor convergence rate. Hence, BCO performance will be improved, allowing for more efficient optimal parameter selection for LSTM.

References

[1] Mrabet, H., Belguith, S., Alhomoud, A., & Jemai, A. (2020). A survey of IoT security based on a layered architecture of sensing and data analysis. *Sensors*, *20*(13), 1-19.

[2] Ullah, F., Naeem, H., Jabbar, S., Khalid, S., Latif, M.A., Al-Turjman, F., & Mostarda, L. (2019). Cyber security threats detection in internet of things using deep learning approach. *IEEE access*, *7*, 124379-124389.

[3] Wani, A., & Revathi, S. (2020). DDoS detection and alleviation in IoT using SDN (SDIoT-DDoS-DA). *Journal of The Institution of Engineers (India): Series B*, *101*(2), 117-128.
[4] Ahamed, M.I., Ahamed, M., Sivaranjani, A., & Chockalingam, S. (2021). Energy bandgap studies on copper chalcogenide semiconductor nanostructures using cohesive energy. *Chalcogenide letters*, *18*(5), 245-253.
[5] Ahamed, M.I., & Kumar, K.S. (2019). Studies on CuSnS quantum dots for O-band wavelength detection. *Materials Science-Poland*, *37*(2), 225-229.
[6] Hu, Z., Wang, L., Qi, L., Li, Y., & Yang, W. (2020). A novel wireless network intrusion detection method based on adaptive synthetic sampling and an improved convolutional neural network. *IEEE Access*, *8*, 195741-195751.
[7] Parra, G. D. L. T., Rad, P., Choo, K. K. R., & Beebe, N. (2020). Detecting Internet of Things attacks using distributed deep learning. *Journal of Network and Computer Applications*, *163*.
[8] Churcher, A., Ullah, R., Ahmad, J., Ur Rehman, S., Masood, F., Gogate, M., & Buchanan, W. J. (2021). An experimental analysis of attack classification using machine learning in IoT networks. *Sensors*, *21*(2), 1-32.
[9] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*(8), 1735-1780.
[10] Niu, B., & Wang, H. (2012). Bacterial colony optimization. *Discrete Dynamics in Nature and Society*, 1-28.
[11] Liang, X., & Znati, T. (2019). A long short-term memory enabled framework for DDoS detection. *In IEEE global communications conference (GLOBECOM)*, 1-6.
[12] Chung, H., & Shin, K.S. (2018). Genetic algorithm-optimized long short-term memory network for stock market prediction. *Sustainability*, *10*(10), 1-18.
[13] Wang, P., Zhao, J., Gao, Y., Sotelo, M.A., & Li, Z. (2020). Lane work-schedule of toll station based on queuing theory and PSO-LSTM model. *IEEE Access*, *8*, 84434-84443.
[14] Ji, Y., Liew, A.W.C., & Yang, L. (2021). A novel improved particle swarm optimization with long-short term memory hybrid model for stock indices forecast. *IEEE Access*, *9*, 23660-23671.
[15] Shao, B., Li, M., Zhao, Y., & Bian, G. (2019). Nickel price forecast based on the LSTM neural network optimized by the improved PSO algorithm. *Mathematical Problems in Engineering*, *2019*, 1-15.
[16] Babu, S.S., & Jayasudha, K. (2022). A Simplex Method-Based Bacterial Colony Optimization for Data Clustering. In *Innovative Data Communication Technologies and Application: Proceedings of ICIDCA*, 987-995. Singapore: Springer Nature Singapore.
[17] Revathi, J., Eswaramurthy, V.P., & Padmavathi, P. (2019). Bacterial colony optimization for data clustering. *In IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, 1-4.
[18] Huang, Y., Gao, Y., Gan, Y., & Ye, M. (2021). A new financial data forecasting model using genetic algorithm and long short-term memory network. *Neurocomputing*, *425*, 207-218.
[19] Koroniotis, N., Moustafa, N., Sitnikova, E., & Turnbull, B. (2019). Towards the development of realistic botnet dataset in the internet of things for network forensic analytics: Bot-iot dataset. *Future Generation Computer Systems*, *100*, 779-796.
[20] Sharafaldin, I., Lashkari, A.H., Hakak, S., & Ghorbani, A.A. (2019). Developing realistic distributed denial of service (DDoS) attack dataset and taxonomy. *In IEEE International Carnahan Conference on Security Technology (ICCST)*, 1-8.
[21] Kalaiselvi, K., Velusamy, K., & Gomathi, C. (2018). Financial prediction using back propagation neural networks with opposition-based learning. *In Journal of Physics: Conference Series*, 1142(1), 1-8. IOP Publishing.
[22] Velusamy, K., & Amalraj, R. (2017). Performance of the cascade correlation neural network for predicting the stock price. *In IEEE Second International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, 1-6.
[23] Ahamed, I., Ahamed, M., Kumar, K.S., & Sivaranjani, A. (2022). Comparative energy bandgap

analysis of zinc and tin based chalcogenide quantum dots. *Revista Mexicana de Física*, *68*(4), 1-8.

[24] Abhishta, A., van Heeswijk, W., Junger, M., Nieuwenhuis, L.J., & Joosten, R. (2020). Why would we get attacked? An analysis of attacker's aims behind DDoS attacks. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA), 11*(2), 3-22.

Authors Biography

Ebtesam Shadadi works as a Lecturer in the Department of Computer Science, College of CS & IT and holds the position of Department Assistant coordinator. Currently pursuing a PhD in the reputed university in London. She is teaching many computer science subjects and having publications and books and giving her contribution in the funded projects. Ebtesam Shadadi holds a Bachelor's degree in Computer Science from King Khalid University, and a Master's degree in Advanced Computer Science from the University of East Anglia, United Kingdom, and she is currently a lecturer in the Department of Computer Science, Jazan University, Saudi Arabia. Her research activity is related to Artificial Intelligence, Machine Learning and Cyber Security.

Latifah Alamer works as a Lecturer in the Department of Information Technology and security, College of CS & IT, Jazan University, KSA and she holds a Bachelor's degree in Information Systems from King Khalid University, and a Master's degree in information Systems from the University of Michigan. USA. Her research activity is related to Artificial Intelligence, Machine Learning and Big data. She Specialized Information Systems Lecturer with abilities in creating and implementing training programs to motivate students. Offering more than 8 years of hands-on experience in teaching at the high education level. Capable of tutoring students and staff on individual basis and ensuring all initiatives are being met. Collaborated with faculties and committees to improve curriculum and instructions for computer-focused courses. Excellent reputation for resolving problems and improving the work process and adapting to new situations and challenges to best enhance the college. Regularly contributed to innovative teaching, research and consultancy developments through workshop and initiatives. She has got American associative of innovation membership 2023- 2027.