Computer Systems
Science and Engineering

Tech Science Press

DOI: 10.32604/csse.2023.037016
Article

Blockchain with Explainable Artificial Intelligence Driven Intrusion Detection for Clustered IoT Driven Ubiquitous Computing System

Reda Salama[1] **and Mahmoud Ragab**[1,2,*]

[1]Information Technology Department, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia
[2]Department of Mathematics, Faculty of Science, Al-Azhar University, Naser City, 11884, Cairo, Egypt
*Corresponding Author: Mahmoud Ragab. Email: mragab@kau.edu.sa

Received: 20 October 2022; Accepted: 06 January 2023

Abstract: In the Internet of Things (IoT) based system, the multi-level client's requirements can be fulfilled by incorporating communication technologies with distributed homogeneous networks called ubiquitous computing systems (UCS). The UCS necessitates heterogeneity, management level, and data transmission for distributed users. Simultaneously, security remains a major issue in the IoT-driven UCS. Besides, energy-limited IoT devices need an effective clustering strategy for optimal energy utilization. The recent developments of explainable artificial intelligence (XAI) concepts can be employed to effectively design intrusion detection systems (IDS) for accomplishing security in UCS. In this view, this study designs a novel Blockchain with Explainable Artificial Intelligence Driven Intrusion Detection for IoT Driven Ubiquitous Computing System (BXAI-IDCUCS) model. The major intention of the BXAI-IDCUCS model is to accomplish energy efficacy and security in the IoT environment. The BXAI-IDCUCS model initially clusters the IoT nodes using an energy-aware duck swarm optimization (EADSO) algorithm to accomplish this. Besides, deep neural network (DNN) is employed for detecting and classifying intrusions in the IoT network. Lastly, blockchain technology is exploited for secure inter-cluster data transmission processes. To ensure the productive performance of the BXAI-IDCUCS model, a comprehensive experimentation study is applied, and the outcomes are assessed under different aspects. The comparison study emphasized the superiority of the BXAI-IDCUCS model over the current state-of-the-art approaches with a packet delivery ratio of 99.29%, a packet loss rate of 0.71%, a throughput of 92.95 Mbps, energy consumption of 0.0891 mJ, a lifetime of 3529 rounds, and accuracy of 99.38%.

Keywords: Blockchain; internet of things; ubiquitous computing; explainable artificial intelligence; clustering; deep learning

1 Introduction

Ubiquitous computing service aims to develop the variety of sensors and networks available to provide timeless services and user location. A major challenge of ubiquitous computing can be context awareness that it can provide numerous services to end users based on potential contextual data. The physical world is converted into a ubiquitous computing environment because of the application and deployment of ubiquitous systems [1]. IoT system service can be developed by incorporating the physical world with computational abilities and decision-making through wireless devices and smart sensing units. Communication networks, Multimedia, business, healthcare, information access, and other applications for commercial and residential customers benefit from this environment [2]. Communication technologies, third-party services, applications, local and distributed resources incorporated with a pervasive computation environment, and querying requests are enhanced by satisfying the users' requirements and providing instant responses. Fig. 1 showcases the general infrastructure of XAI.

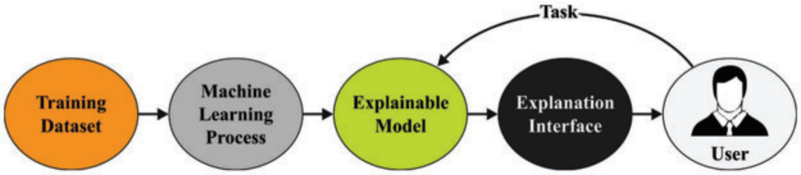


Figure 1: General structure of XAI

Cloud and other network services are fetched by a layered technique in this environment to guarantee the IoT-enabled wearable sensor has sufficient access to resources at the user network edge [3,4]. A clustering method can partition sensors into distinct groups or clusters. Different security strategies and technologies have been developed to protect network security. Intrusion detection is a traditional network security technique [5–7]. An earlier intrusion detection system (IDS) mainly utilizes misuse detection. Misuse detection records the attack through a signature database and judges an intrusion with the data or event matching the signature. But misuse detection is non-practical since it could identify unrecorded attacks. The anomaly detection technique is currently commonly utilized [8] with machine learning (ML) advancement.

Explainable artificial intelligence (XAI) is a group of methods and processes that enables users to trust and understand the output and results generated by ML algorithms [9]. Explainable AI describes an AI technique, potential biases, and expected impact [10]. It can describe model outcomes, accuracy, fairness, and transparency in AI-assisted decision-making. Recently, the development of the blockchain (BC) technique has mentioned a path worth attempting to resolve the distributed trust issue in the IoT platform [11]. BC is a peer-to-peer distributed system with decentralization, non-tempera, system autonomy, and transparency [12], which could successfully improve network collaboration and device security.

This study designs a novel Blockchain with Explainable Artificial Intelligence Driven Intrusion Detection for IoT Driven Ubiquitous Computing System (BXAI-IDCUCS) model. The proposed BXAI-IDCUCS model initially clusters the IoT nodes using the energy-aware duck swarm optimization (EADSO) algorithm. Besides, deep neural network (DNN) is employed for detecting and classifying intrusions in the IoT network. Blockchain (BC) technology is exploited for secure inter-cluster data transmission processes. A comprehensive experimentation study is performed to ensure the productive performance of the BXAI-IDCUCS model, and the results are assessed under several aspects.

2 Related Works

Liu et al. [13] presented a cooperative intrusion detection (ID) system that offloads the trained models for distributing edge devices (for instance, related to vehicle and roadside units (RSUs). The distributing federated-based model reduces the utilization of resources of the centralized servers; however, privacy and security are assured. BC was utilized to store and share the trained methods for ensuring the security of the aggregation method. The data-driven trust process dependent upon blockchain was projected as a decentralized and energy-effectual solution to detect internal attacks from IoT-driven SN [14]. During the grey and black hole attack setting, the message overhead was enhanced utilizing the presented technique related to the present solution. In both grey and black hole attacks, the time obtained to detect malicious nodes also decreased significantly.

In [15], a block-chain-based authentication method was presented to secure routing from the WSNs. The malicious and unauthenticated nodes affected the routing procedure, and the accurate detection of routing direction developed a challenging problem. Thus, during this method, the registration of nodes was completed by Certificate Authority Node (CAN) to prevent the contribution of malicious nodes from the networks. The authors in [16] examined the possible threat in SDN-empowered WSN and detailed black hole attacks. During the case, it can also be presented a new lightweight security method which exploits the BC method that capable of protecting the flow tables from all the nodes, which is an essential target of a feasible routing attack. An unchangeable fingerprint named the signature token to the flow entry can be created with a secret key going to all the nodes.

Mahapatra et al. [17] presented a Quantum Atom Search Optimization with BC-assisted Data Transmission (QASO-BDT) method to a relay node election with security-supported data transmissions. This method contains 3 stages, namely transmission, registration, and clustering. Primarily, under the node registration stage, all sensor nodes (SNs) obtain registration from the blockchain network with Capillary Gateway (CG). Afterwards, a CH was chosen under the clustering stage, and an improved multi-view clustering method was utilized for clustering the node into distinct clusters. Wang et al. [18] utilized SHapley Additive exPlanations (SHAP) and integrated local and global details to improve the interpretation of IDS. The local explanation explains why the method creates particular decisions based on the particular input. The authors in [19] progressed a new secure unequal clustering protocol with an ID approach for achieving QoS parameters such as security, energy, and lifetime.

3 The Proposed Model

The BXAI-IDCUCS model has been developed in this study to accomplish maximum energy efficacy and security in the IoT environment. The BXAI-IDCUCS model aims to find the existence of intrusions in the clustered IoT environment and perform blockchain-driven secure data transmission. The BXAI-IDCUCS model follows a three-stage process: clustering, intrusion detection, and BC-based data transmission. Fig. 2 illustrates the workflow of the BXAI-IDCUCS algorithm.

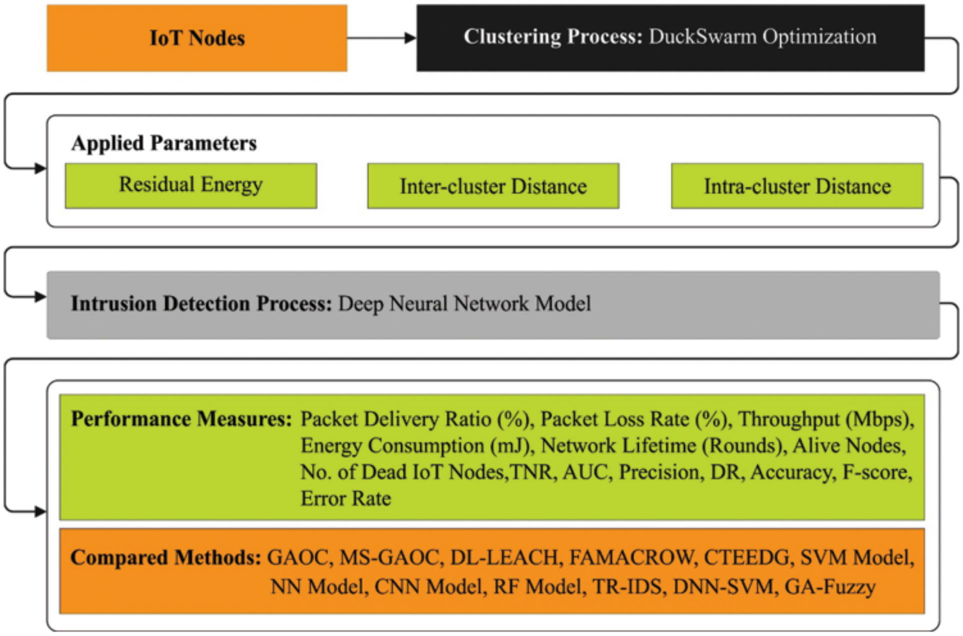

IoT Nodes
Clustering Process: DuckSwarm Optimization
Applied Parameters
Residual Energy
Inter-cluster Distance
Intra-cluster Distance
Intrusion Detection Process: Deep Neural Network Model
Performance Measures: Packet Delivery Ratio (%), Packet Loss Rate (%), Throughput (Mbps), Energy Consumption (mJ), Network Lifetime (Rounds), Alive Nodes, No. of Dead IoT Nodes,TNR, AUC, Precision, DR, Accuracy, F-score, Error Rate
Compared Methods: GAOC, MS-GAOC, DL-LEACH, FAMACROW, CTEEDG, SVM Model, NN Model, CNN Model, RF Model, TR-IDS, DNN-SVM, GA-Fuzzy

Figure 2: Workflow of BXAI-IDCUCS technique

3.1 EADSO-based Clustering Technique

The EADSO algorithm with a fitness function involving three variables is employed for the clustering process. The EADSO algorithm is derived from the behaviour of ducks. Three major processes of the DSO algorithm are given in the following: position of duck swarm afterwards queuing (Population initialization), search for a food source (Exploration stage), and foraging in the group (Exploitation stage). Note that two rules need to be obeyed in searching for food for ducks [20].

- While searching for food, ducks with stronger searching capacity are situated near the centre of the food source, which attracts other individuals to get close to them; the upgraded position is affected by neighbouring individuals.
- During foraging, all the individuals approach the food; the following location is affected by nearby individuals and the leader duck or food position.

The randomly generated initial location in the D-dimension searching space can be expressed by:

$$X_i = L_b + (U_b - L_b) \cdot 0 \tag{1}$$

Whereas X_i signifies the spatial location of i-th duck ($i = 1, 2, 3, N$), N denotes the amount of population size. L_b and U_b represent the upper and lower limits, and 0 denotes an arbitrary value within $(0, 1)$. Afterwards, the duck's queuing behaviour arrived at a position with more food. All the individuals disperse gradually and start to search for food; this procedure is described in the following:

$$X_i^{t+1} = \begin{cases} X_i^t + \mu \cdot X_i^t \cdot sign(r - 0.5), P > rand \\ X_i^t + CF_1 \cdot (X_{leader}^t - X_i^t) + CF_2 \cdot (X_j^t - X_i^t), P < rand \end{cases} \tag{2}$$

In which F_1 and CF_2 denote cooperation and competition coefficients among ducks in the searching process, correspondingly, $sign(r - 0.5)$ denotes an effect on the food searching process. It is fixed to -1 or 1. μ signifies the control variable of global searching, P denotes the search conversion possibility of the exploration stage, and $C\ X^t_{leader}$ characterizes the optimal duck location of the current value in t-the iteration. x^t_j signifies the agent around X^t_i looking for food by a group of ducks in the t-th iteration. Furthermore, variable μ is evaluated by:

$$\mu = K \cdot \left(1 - \frac{t}{t_{max}}\right) \tag{3}$$

Now, K is evaluated as follows:

$$K = \sin\left(2 \cdot rand\right) + 1 \tag{4}$$

After the food searching process, sufficient food satisfies the foraging of the duck. This procedure is strongly associated with the fitness of duck location, and it is described in the following:

$$X^{t+1}_i = \begin{cases} X^t_i + \mu \cdot (X^t_{leader} - X^t_i), f(X^t_i) > f(X^{t+1}_i) \\ X^t_i + K_1 \cdot (X^t_{leader} - X^t_i) + KF_2 \cdot (X^t_k - X^t_j), else \end{cases} \tag{5}$$

In which KF_1 and KF_2 parameters signify the cooperation and competition coefficients among ducks in the exploitation stage, correspondingly μ means the control variables of global searching in the exploitation stage, X^t_k and X^t_j represent the agent around X^t_i in foraging of the duck group in the t-th iteration, X^t_{leader} characterizes the optimal duck location of the existing value in the t-th iteration. Whereas $k \neq j$.

Note that the parameter values CF_1, CF_2, KF_1 and KF_2 are within $(0, 2)$, and the following equation can compute it:

$$CF_i\ or\ KF_i \leftarrow \frac{1}{FP} \cdot rand(0, 1)(i = 1, 2) \tag{6}$$

where FP denotes constant, it is fixed to 0.618; the rand indicates an arbitrary value within $(0, 1)$.

In the exploitation stage, the procedure of duck upgrades its location pertaining to $X_i, X_j\cdot, X_k$ and X_{eader} in a 2D searching space. Path 1 represents the selection of ducks with cooperation. Path 2 characterizes the competition *between* X_i *and* X_k *and* X_j in the t-th iteration. Path 3 denotes the selection of the duck that fails to compete. The EADSO algorithm includes fitness variables such as residual energy (RE), inter-cluster distance, and intra-cluster distance.

Residual Energy

CH performs several activities: data communication, gathering, sensing, aggregation, and so on. Consequently, CH intakes the highest energy compared to others. Then, it is essential to describe an FF that shares the loads between every sensor in the network. The following equations show the fitness parameter utilized for effective network usage.

$$R_e = e\left(n_i\right)$$

$$Avg_e = \frac{1}{n}\sum_{i=0}^{n} e\left(n_i\right)$$

$$f_1 = CH_{opt} * \frac{R_e}{Avg_e} = \frac{CH_{opt} * e\left(n_i\right)}{\frac{1}{n}\sum_{i=0}^{n} e\left(n_i\right)} \forall CH_{opt} = 5\%\ of\ n,\ e\left(n_i\right)$$
$$= 0.5\ J\ or 1.25\ J\ or\ 1.75\ J \tag{7}$$

R_e, Avg_e, & n_i represent the network's residual node energy, normal energy, and the total number of sensors in the network. CH_{opt} specifies the optimum percentage of CH. A value of f_1 demonstrates the ratio of $Avg_e$and R_e.

Intra-Cluster Distance: SNs transfer the information to CH. When the CHs are farther from CM, then the sensor depletes energy. In case when CHs are nearer to the member sensor, it employs the lowest energy.

$$f_2 = \frac{1}{n_{s\tau}} \sum_{i=0}^{n_{sr}} disT(CH, i) \forall dist(CH, i) = 1 \ to \ 35\,m, n_{sr} = 1 \ to \ 100 \tag{8}$$

Here, n_{sr} & $dist(CH, i)$ indicates the number of SNs and Euclidean distance from CH and node in the sensing sequence. Therefore, the values of f_3 should be lesser while reducing the intracluster transmission energy.

Inter-Cluster Distance: While performing CH selection, the distance between BS and CH is essential. When the chosen CH is farther from the sink, it employs energy quickly and is assessed in the equation,

$$f_3 = \frac{1}{CH} \sum_{\mathrm{i=0}}^{\mathrm{CH}} dist(BS, CH_i) \forall dist(BS, CH_i) = 1 \ to \ 70m, CH = 1 \ to \ 15 \tag{9}$$

where $dist(BS, CH_i)$ represents the Euclidean distance between BS and CH_i. The value of f_3 is minimized, meaning the selected CH is not far from the BS.

When f_1,f_2, $and f_3$ are calculated, the cost function named FF can be defined,

$$F = Maximize\ Fitness = \alpha * f_1 + \beta * \frac{1}{f_2} + \gamma * \frac{1}{f_3} \tag{10}$$

Let α, β, γ be the weight coefficients of f_1,f_2,f_3, and FF variables. The weight coefficient ranges from zero to one.

3.2 Intrusion Detection Process

In this phase, the DNN model is employed to detect and classify intrusions in the IoT system. The DNN method comprised hidden, input, and resultant layers. In the training phase, DNN increases the node weight in the hidden layer [21]. Due to the gradual increase in training iteration, the NN often fits the labelled training information solution boundary. DNN, classifier accuracy, and 2 hidden layers were introduced to increase the training method's speed. In the hidden layer, entire nodes are described in the following.

$$n = \sqrt{a + b} + c \tag{11}$$

The number of input layers is characterized by a, and the amount of resultant layers is represented as b, the amount of hidden layers is symbolized as n, and a constant value from [1, 10] is indicated as c.

$$S = \frac{1}{1 + e^{-x}} \tag{12}$$

The input dataset of the system is called x, which can be assisted by a mapping function M_f.

$$M_f = sigm\,(\omega_i x + \beta_i) \tag{13}$$

Ω and β denote the weight matrix and a bias between resultant and hidden layers. As well it is developed by labelled data samples (x, l) for the hidden layer, and a loss structure can be defined by,

$$S\left(W_s, b_s; x, l\right) = \frac{1}{2m} \sum_{j=1}^{m} \left\| h_j\left(W_s, b_s; x\right) - l_j \right\|_2^2 \tag{14}$$

Now W_s and b_s determine a subset of bias, and m represents the number of neurons in a hidden layer.

Cross entropy (CE) is employed as a loss function of DNN, regarded as the testing and training configuration. The application of CE doesn't employ the function of the sigmoid and softmax output framework. It can be expressed as follows

$$C_E = \frac{1}{n} \sum_{k=1}^{n} \left[Y_k log \hat{Y}_k + (1 - Y_k) \log(1 - \hat{Y}_k) \right] \tag{15}$$

where n denotes the amount of training sample, Y_k denotes the kth original training set results, and $\hat{Y}_k$ represents the kth determined testing set results.

3.3 Blockchain-Driven Secure Data Transmission

In this work, blockchain technology is exploited for secure inter-cluster data transmission. Generally, BC meant that group of blocks. In these blocks, a single block comprises 4 segmented data concerning the transaction (Bitcoin, Ethereum), Hash value of the existing block, Timestamp, and Previous block [22]. In addition, the BC was determined as distributed, and the usual digital ledger was utilized to save the transaction data in the diverse point. Therefore, when an attacker tries to derive information, it can be difficult as all blocks have the cryptographic value of preceding blocks. Fig. 3 defines the framework of BC. At this point, every transaction was reached in the application of cryptographic hash value verified by all the miners. It can be taken with the same value as a completed ledger and contains blocks of every transaction. The decentralized saved is another source from BC, and a superior count of data was saved and connected in the existing block for the preceding block utilizing smart contract code. Swarm, SiacoinDB, BigchainDB, LitecoinDB, MoneroDB, Interplanetary File System (IPFS), and several other factors are presently executed to the decentralized database.

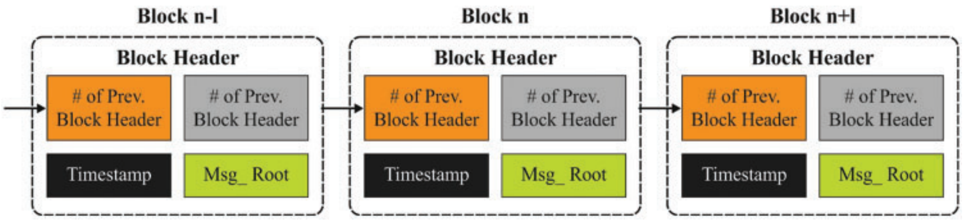


Figure 3: Structure of blockchain

4 Performance Validation

In this section, a detailed validation of the BXAI-IDCUCS model is carried out under distinct IoT nodes. Table 1 provides a comprehensive PDR and PLR examination of the BXAI-IDCUCS method with recent models. Fig. 4 portrays a close PDR inspection of the BXAI-IDCUCS technique under different IoT nodes. The results implied that the BXAI-IDCUCS approach had gained maximum PDR values over the other models. On 100 IoT nodes, the BXAI-IDCUCS technique has accomplished a maximum PDR of 98.71%, whereas GAOC, MS-GAOC, DL-Leach, FAMACROW, and CTEEDG models have obtained reduced PDR of 93.95%, 93.37%, 94.64%, 97.55%, and 97.84% correspondingly. Furthermore, on 500 IoT nodes, the BXAI-IDCUCS approach has reached a high PDR of 98.95%, whereas GAOC, MS-GAOC, DL-Leach, FAMACROW, and CTEEDG models have obtained lesser PDR of 80.55%, 83.28%, 81.92%, 87.75%, and 88.75% correspondingly.

Table 1: PDR and PLR analysis of BXAI-IDCUCS technique under dissimilar IoT nodes

No. of IoT nodes	GAOC	MS-GAOC	DL-Leach	FAMACROW	CTEEDG	BXAI-IDCUCS
Packet delivery ratio (%)						
100	93.95	93.37	94.64	97.55	97.84	98.71
200	90.07	91.03	92.33	94.04	95.96	99.29
300	86.94	88.28	88.35	91.97	92.30	98.55
400	83.24	84.83	85.48	90.12	90.59	99.04
500	80.55	83.28	81.92	87.75	88.75	98.95
Packet loss rate (%)						
100	6.05	6.63	5.36	2.45	2.16	1.29
200	9.93	8.97	7.67	5.96	4.04	0.71
300	13.06	11.72	11.65	8.03	7.70	1.45
400	16.76	15.17	14.52	9.88	9.41	0.96
500	19.45	16.72	18.08	12.25	11.25	1.05

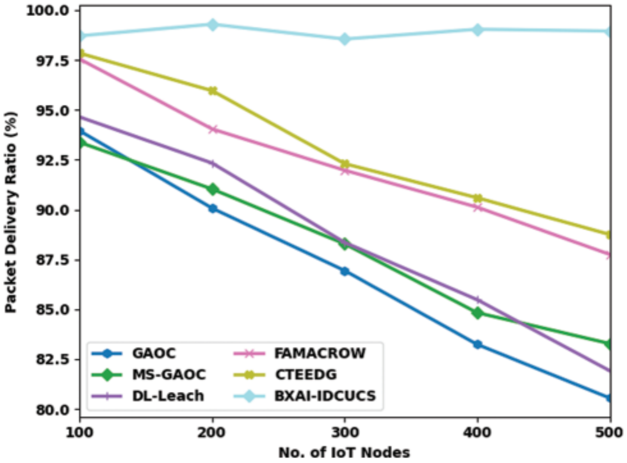


Figure 4: PDR analysis of BXAI-IDCUCS technique under dissimilar IoT nodes

A detailed PLR assessment of the BXAI-IDCUCS model with existing approaches is performed in Fig. 5. The outcome shows that the BXAI-IDCUCS model has gained effectual outcomes with minimal values of PLR. On 100 IoT nodes, the BXAI-IDCUCS model has offered a reduced PLR of

1.29%, whereas GAOC, MS-GAOC, DL-Leach, FAMACROW, and CTEEDG models have gained higher PLR of 6.05%, 6.63%, 5.36%, 2.45%, and 2.16% respectively. Furthermore, on 500 IoT nodes, the BXAI-IDCUCS model has an accessible reduced PLR of 1.05%. In contrast, GAOC, MS-GAOC, DL-Leach, FAMACROW, and CTEEDG models have reached superior PLRs of 19.45%, 16.72%, 18.08%, 12.25%, and 11.25% correspondingly.

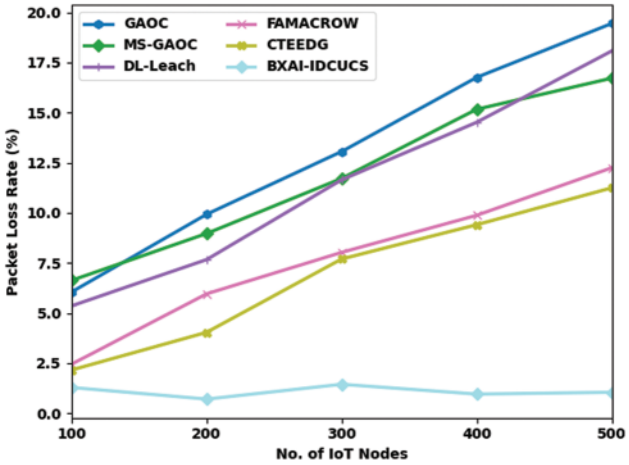


Figure 5: PLR analysis of BXAI-IDCUCS technique under dissimilar IoT nodes

Table 2 and Fig. 6 demonstrate a comparative throughput (THPT) analysis of the BXAI-IDCUCS method under different IoT nodes. The results implied that the BXAI-IDCUCS approach had gained maximal THPT values over the other methods. On 100 IoT nodes, the BXAI-IDCUCS system has achieved enhanced THPT of 92.95 Mbps whereas GAOC, MS-GAOC, DL-Leach, FAMACROW, and CTEEDG models have obtained lower THPT of 67.57, 68.69, 71.7, 76.59, and 82.44 Mbps correspondingly. In addition, on 500 IoT nodes, the BXAI-IDCUCS technique has achieved a maximum THPT of 81.88 Mbps. In contrast, GAOC, MS-GAOC, DL-Leach, FAMACROW, and CTEEDG techniques have reduced THPT of 53.58 and 56.48 Mbps, respectively 61.84, 61.38, and 72.97 Mbps correspondingly.

Table 2: Throughput analysis of BXAI-IDCUCS technique under dissimilar IoT nodes

Throughput (Mbps)						
No. of IoT nodes	GAOC	MS-GAOC	DL-Leach	FAMACROW	CTEEDG	BXAI-IDCUCS
100	67.57	68.69	71.7	76.59	82.44	92.95
200	64.08	66.75	69.29	72.74	78.72	90.78
300	60.31	63.79	66.06	68.83	77.05	86.96
400	56.58	60.06	63.77	65.19	74.93	84.44
500	53.58	56.48	61.84	61.38	72.97	81.88

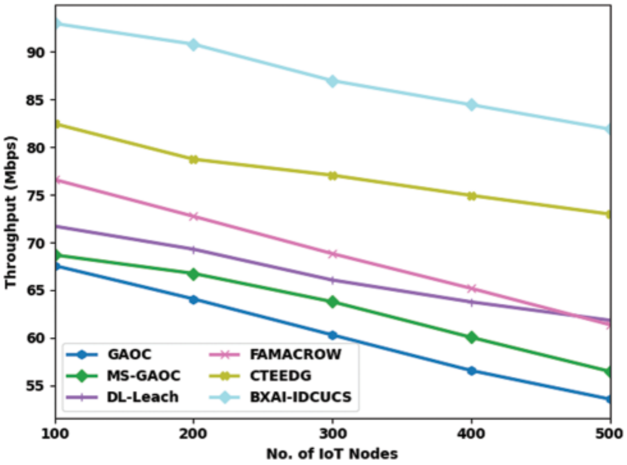


Figure 6: Throughput analysis of BXAI-IDCUCS technique under dissimilar IoT nodes

With recent approaches, a brief ECM assessment of the BXAI-IDCUCS technique is performed in Table 3 and Fig. 7. The experimental outcome indicated that the BXAI-IDCUCS system had gained effectual outcomes with minimal values of ECM. On 100 IoT nodes, the BXAI-IDCUCS system has accessible lower ECM of 0.0891 mJ whereas GAOC, MS-GAOC, DL-Leach, FAMACROW, and CTEEDG methodologies have gained maximal ECM of 0.2234, 0.2044, 0.1753, 0.1566, and 0.1183 mJ respectively. Besides, on 500 IoT nodes, the BXAI-IDCUCS model has offered a reduced ECM of 0.3257 mJ whereas GAOC, MS-GAOC, DL-Leach, FAMACROW, and CTEEDG models have gained maximum ECM of 0.4391, 0.4219, 0.3964, 0.3798, and 0.2825 mJ correspondingly.

Table 3: Energy consumption analysis of BXAI-IDCUCS technique under dissimilar IoT nodes

Energy consumption (mJ)						
No. of IoT nodes	GAOC	MS-GAOC	DL-Leach	FAMACROW	CTEEDG	BXAI-IDCUCS
100	0.2234	0.2044	0.1753	0.1566	0.1183	0.0891
200	0.2924	0.2428	0.2339	0.2001	0.1584	0.1380
300	0.3571	0.3113	0.3000	0.2612	0.2026	0.2057
400	0.3954	0.3686	0.3531	0.3182	0.2439	0.2748
500	0.4391	0.4219	0.3964	0.3798	0.2825	0.3257

Table 4 and Fig. 8 illustrate a comparative NLT analysis of the BXAI-IDCUCS system under different IoT nodes. The results exposed that the BXAI-IDCUCS model has gained maximum NLT values over the other models. On 100 IoT nodes, the BXAI-IDCUCS model has achieved higherNLT of 2025 rounds, whereas GAOC, MS-GAOC, DL-Leach, FAMACROW, and CTEEDG models have obtained reduced NLT of 1318, 1366, 1305, 1351, and 1713 rounds correspondingly. Eventually, on 500 IoT nodes, the BXAI-IDCUCS methodology has achieved a maximum NLT of 3529 rounds. In contrast, GAOC, MS-GAOC, DL-Leach, FAMACROW, and CTEEDG approaches have obtained reduced NLT of 2483, 2308, 2537, 2670, and 3354 rounds correspondingly.

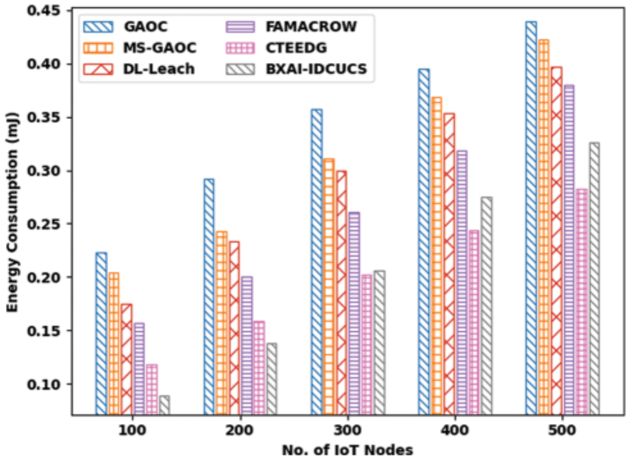


Figure 7: ECM analysis of BXAI-IDCUCS technique under dissimilar IoT nodes

Table 4: Network lifetime analysis of BXAI-IDCUCS technique under dissimilar IoT nodes

Network lifetime (Rounds)						
No. of IoT nodes	GAOC	MS-GAOC	DL-Leach	FAMACROW	CTEEDG	BXAI-IDCUCS
100	1318	1366	1305	1351	1713	2025
200	1765	1552	1618	1702	2159	2416
300	1976	1705	1837	2009	2539	2866
400	2303	1864	2099	2284	2923	3255
500	2483	2308	2537	2670	3354	3529

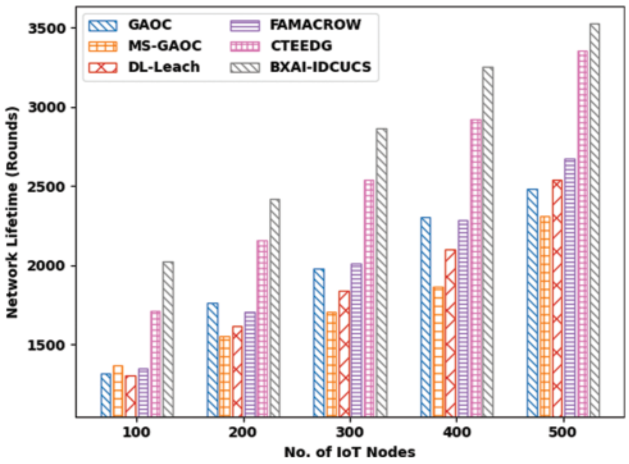
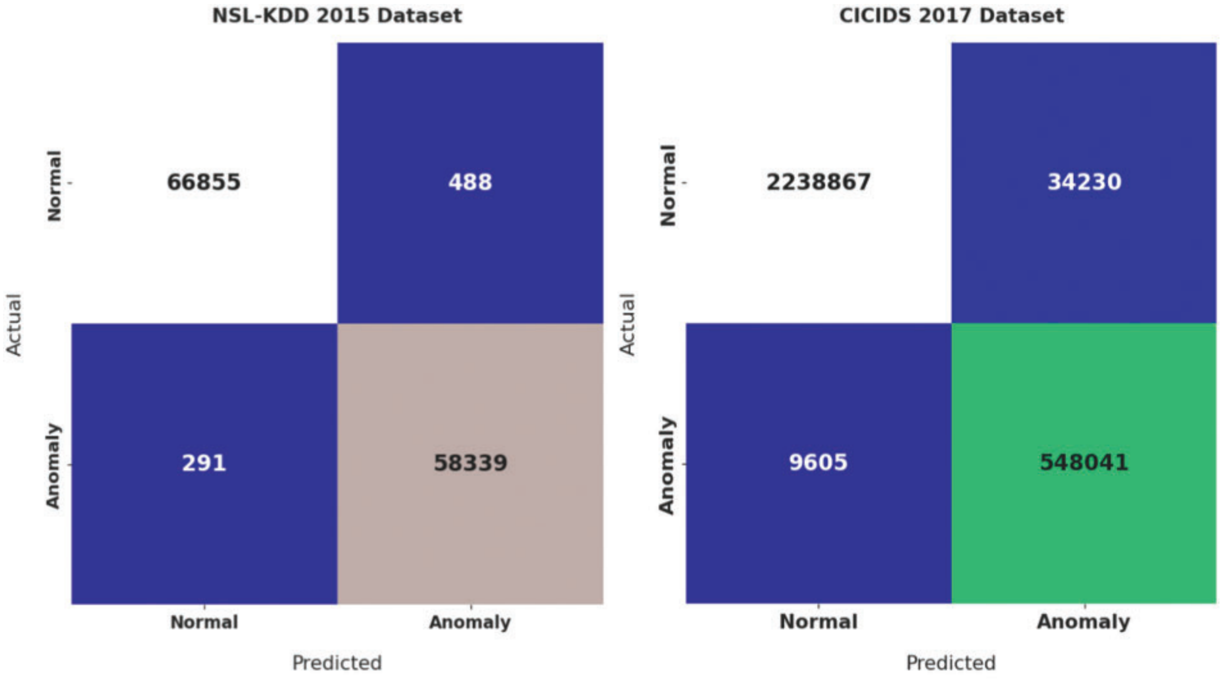


Figure 8: NLT analysis of BXAI-IDCUCS technique under dissimilar IoT nodes

Next, the performance of intrusion detection of the BXAI-IDCUCS model is validated using two benchmark datasets, as given in Table 5. Fig. 9 portrays a pair of confusion matrices offered by the BXAI-IDCUCS model on the test dataset. On NSL-KDD 2015 dataset, the BXAI-IDCUCS model has identified 66855 samples under the normal class and 291 samples under the anomaly class. Besides, on CICIDS 2017 dataset, the BXAI-IDCUCS approach has identified 2238867 samples under the normal class and 548041 samples under the anomaly class.

Table 5: Dataset description

Dataset	No. of instances	No. of attributes	No. of classes	Normal/Anomaly
NSL-KDD 2015	125973	41	2	67343/58630
CICIDS 2017	2830743	80	2	2273097/557646



(a) (b)

Figure 9: Confusion matrix of BXAI-IDCUCS technique under two datasets

Table 6 reports a brief IDS outcome of the BXAI-IDCUCS model on two datasets. With the NSL-KDD2015 dataset, the BXAI-IDCUCS model has offered average accuracy of 99.38%, precision of 99.37%, DR of 99.39%, TNR of 99.39%, F-score of 99.38%, AUC of 99.39%, and the error rate of 0.62. Moreover, with CICIDS 2017 dataset, the BXAI-IDCUCS system has obtainable average $accu_y$ of 98.53%, $prec_n$ of 97.11%, DR of 98.35%, TNR of 98.35%, F_{score} of 97.71%, AUC of 98.35%, and the error rate of 1.47.

Table 6: Result analysis of the BXAI-IDCUCS technique with different measures

Class Labels	$Accu_y$	$Prec_n$	DR	TNR	F_{score}	AUC	Error Rate
NSL-KDD 2015							
Normal	99.38	99.57	99.28	99.5	99.42	99.39	0.62
Anomaly	99.38	99.17	99.5	99.28	99.34	99.39	0.62
Average	**99.38**	**99.37**	**99.39**	**99.39**	**99.38**	**99.39**	**0.62**
CICIDS 2017							
Normal	98.53	99.51	98.65	98.04	99.08	98.35	1.47
Anomaly	98.53	94.71	98.04	98.65	96.35	98.35	1.47
Average	**98.53**	**97.11**	**98.35**	**98.35**	**97.71**	**98.35**	**1.47**

A brief comparative study of the BXAI-IDCUCS with recent models is made in Table 7 [23–26]. Fig. 10 inspects a comparative accuracy examination of the BXAI-IDCUCS with recent models. The figure indicated that the SVM system had offered a lower $accu_y$ of 87.16%. Followed by the NN and DNN-SVM approaches have obtained somewhat enhanced $accu_y$ of 90.99% and 92.03%, correspondingly. In line with this, the GA-Fuzzy and CNN models have correspondingly resulted in $accu_y$ of 96.53% and 96.75%. The RF and TR-IDS models have also accomplished reasonable $accu_y$ of 98.21% and 99.10%. But the BXAI-IDCUCS model has obtained the highest $accu_y$ of 99.38%.

Table 7: Comparative analysis of BXAI-IDCUCS technique with recent algorithms

Methods	$Accu_y$	DR
SVM Model	87.16	80.48
NN Model	90.99	92.17
CNN Model	96.75	97.61
RF Model	98.21	97.81
TR-IDS	99.10	99.25
DNN-SVM	92.03	95.32
GA-Fuzzy	96.53	97.38
BXAI-IDCUCS	99.38	99.39

Fig. 11 demonstrates a comparative DR examination of the BXAI-IDCUCS approach with recent models. The figure revealed that the SVM method offered a lower DR of 80.48%. Followed by the NN and DNN-SVM methods have obtained somewhat higher DR of 92.17% and 95.32%, correspondingly. Also, the GA-Fuzzy and CNN models have resulted in DR of 97.38% and 97.61%, correspondingly. Along with that, the RF and TR-IDS techniques have accomplished reasonable DR of 97.81% and 99.25%. But the BXAI-IDCUCS method has gained maximum DR of 99.39%. Afterward inspecting the results and discussion, it is confirmed that the BXAI-IDCUCS model has accomplished maximum energy efficacy and security.

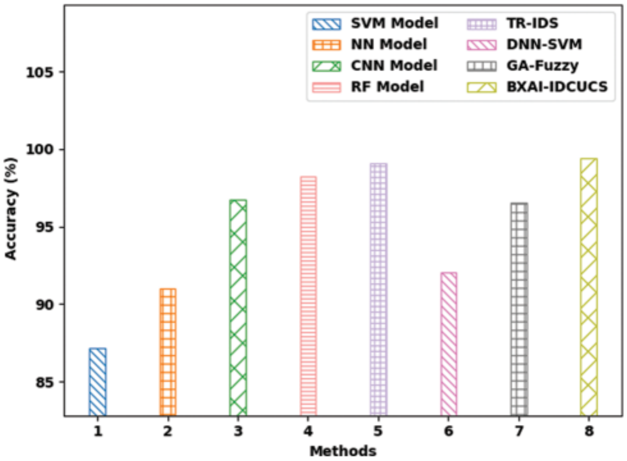


Figure 10: Accuracy analysis of BXAI-IDCUCS technique with recent algorithms

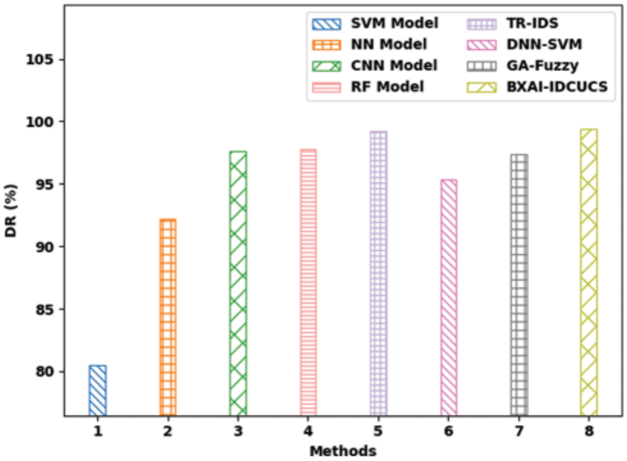


Figure 11: DR analysis of BXAI-IDCUCS technique with recent algorithms

5 Conclusion

In this study, the BXAI-IDCUCS model has been developed to accomplish maximum energy efficacy and security in the IoT environment. The BXAI-IDCUCS model follows a three stage process namely clustering, intrusion detection, and blockchain based data transmission. For clustering process, the EADSO algorithm with fitness function involving three variables is employed. In addition, DNN model was utilized for the detection and classification of intrusions that exist in the IoT network. Lastly, BC technology is exploited for secure inter-cluster data transmission process. To assure effectual performance of the BXAI-IDCUCS model, a comprehensive experimentation study is applied and the outcomes are assessed under several aspects. The comparison study highlighted the superiority of the BXAI-IDCUCS approach over the recent state of art approaches with packet delivery ratio of 99.29%, packet loss rate of 0.71%, throughput of 92.95 Mbps, energy consumption of 0.0891 mJ, lifetime of 3529 rounds, and accuracy of 99.38%. In the future, multihop route selection models can be developed for optimal load balancing in the IoT environment.

Funding Statement: This research work was funded by Institutional Fund Projects under grant no. (IFPIP: 624-611-1443). Therefore, the authors gratefully acknowledge technical and financial support provided by the Ministry of Education and Deanship of Scientific Research (DSR), King Abdulaziz University (KAU), Jeddah, Saudi Arabia.

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

References

[1] Y. Maleh, Y. Baddi, M. Alazab, L. Tawalbeh and I. eds Romdhani, "Artificial intelligence and blockchain for future cybersecurity applications," in *Studies in Big Data*, vol. 90, Cham: Springer, 2021.

[2] D. Sivaganesan, "A data driven trust mechanism based on blockchain in IoT sensor networks for detectio., mitigation of attacks," *Journal of Trends in Computer Science and Smart Technology*, vol. 3, no. 1, pp. 59–69, 2021.

[3] S. Arjunan and S. Pothula, "A survey on unequal clustering protocols in wireless sensor networks," *Journal of King Saud University-Computer and Information Sciences*, vol. 31, no. 3, pp. 304–317, 2019.

[4] S. Awan, M. B. E. Sajid, S. Amjad, U. Aziz, U. Gurmani *et al.,* "Blockchain based authentication and trust evaluation mechanism for secure routing in wireless sensor networks," in *Int. Conf. on Innovative Mobile and Internet Services in Ubiquitous Computing, Lecture Notes in Networks and Systems book series*, Cham, Springer, vol. 279, pp. 96–107, 2021.

[5] R. Goyat, G. Kumar, M. Alazab, R. Saha, R. Thomas *et al.,* "A secure localization scheme based on trust assessment for WSNs using blockchain technology," *Future Generation Computer Systems*, vol. 125, no. 12, pp. 221–231, 2021.

[6] R. Mansour, "Artificial intelligence based optimization with deep learning model for blockchain enabled intrusion detection in CPS environment," *Scientific Reports*, vol. 12, no. 1, pp. 1–14, 2022.

[7] M. Ragab and M. F. Sabir, "Outlier detection with optimal hybrid deep learning enabled intrusion detection system for ubiquitous and smart environment," *Sustainable Energy Technologies and Assessments*, vol. 52, no. 1, pp. 102311, 2022.

[8] S. Arjunan and P. Sujatha, "Lifetime maximization of wireless sensor network using fuzzy based unequal clustering and ACO based routing hybrid protocol," *Applied Intelligence*, vol. 48, no. 8, pp. 2229–2246, 2018.

[9] A. A. Anitha and L. Arockiam, "A review on intrusion detection systems to secure IoT networks," *International Journal of Computer Networks and Applications*, vol. 9, no. 1, pp. 38–50, 2022.

[10] S. Arjunan, S. Pothula and D. Ponnurangam, "F5N-based unequal clustering protocol (F5NUCP) for wireless sensor networks," *International Journal of Communication Systems*, vol. 31, no. 17, pp. e3811, 2018.

[11] V. Chang, L. Golightly, P. Modesti, Q. A. Xu, L. M. T. Doan *et al.,* "A survey on intrusion detection systems for fog and cloud computing," *Future Internet*, vol. 14, no. 3, pp. 89, 2022.

[12] U. Farooq, N. Tariq, M. Asim, T. Baker and A. Al-Shamma'a, "Machine learning and the internet of things security: Solutions and open challenges," *Journal of Parallel and Distributed Computing*, vol. 162, no. 1, pp. 89–104, 2022.

[13] H. Liu, S. Zhang, P. Zhang, X. Zhou, X. Shao *et al.,* "Blockchain and federated learning for collaborative intrusion detection in vehicular edge computing," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 6, pp. 6073–6084, 2021.

[14] D. Sivaganesan, "A data driven trust mechanism based on blockchain in IoT sensor networks for detectio., mitigation of attacks," *Journal of trends in Computer Science and Smart technology (TCSST)*, vol. 3, no. 1, pp. 59–69, 2021.

[15] U. Aziz, M. U. Gurmani, S. Awan, M. B. E. Sajid, S. Amjad *et al.,* "A blockchain based secure authentication and routing mechanism for wireless sensor networks," in *Int. Conf. on Innovative Mobile*

and Internet Services in Ubiquitous Computing, Lecture Notes in Networks and Systems Book Series, Cham, Springer, vol. 279, pp. 87–95, 2021.

[16] E. Karakoç and C. Çeken, "Black hole attack prevention scheme using a blockchain-block approach in SDN-enabled WSN," *International Journal of Ad Hoc and Ubiquitous Computing*, vol. 37, no. 1, pp. 37–49, 2021.

[17] S. N. Mahapatra, B. K. Singh and V. Kumar, "A secure multi-hop relay node selection scheme based data transmission in wireless ad-hoc network via block chain," *Multimedia Tools and Applications*, vol. 81, no. 13, pp. 18343–183737, 2022.

[18] M. Wang, K. Zheng, Y. Yang and X. Wang, "An explainable machine learning framework for intrusion detection systems," *IEEE Access*, vol. 8, pp. 73127–73141, 2020.

[19] M. Maheswari and R. A. Karthika, "A novel QoS based secure unequal clustering protocol with intrusion detection system in wireless sensor networks," *Wireless Personal Communications*, vol. 118, no. 2, pp. 1535–1557, 2021.

[20] M. Zhang, G. Wen and J. Yang, "Duck swarm algorithm: A novel swarm intelligence algorithm," arXiv:2112.13508, 2021.

[21] P. Devan and N. Khare, "An efficient XGBoost-DNN-based classification model for network intrusion detection system," *Neural Computing and Applications*, vol. 32, no. 16, pp. 12499–12514, 2020.

[22] G. N. Nguyen, N. H. Le Viet, A. F. S. Devaraj, R. Gobi and K. Shankar, "Blockchain enabled energy efficient red deer algorithm based clustering protocol for pervasive wireless sensor networks," *Sustainable Computing: Informatics and Systems*, vol. 28, pp. 100464, 2020.

[23] E. Min, J. Long, Q. Liu, J. Cui and W. Chen, "TR-IDS: Anomaly-based intrusion detection through text-convolutional neural network and random forest," *Security and Communication Networks*, vol. 2018, no. 1, pp. 1–9, 2018.

[24] M. Ragab and A. Altalbe, "A blockchain-based architecture for enabling cybersecurity in the internet-of-critical infrastructures," *CMC-Computers Materials & Continua*, vol. 72, pp. 1579–1592, 2022.

[25] R. Mansour, "Blockchain assisted clustering with intrusion detection system for industrial internet of things environment," *Expert Systems with Applications*, vol. 207, no. 14, pp. 117995, 2022.

[26] K. Karunanithy and B. Velusamy, "Cluster-tree based energy efficient data gathering protocol for industrial automation using WSNs and IoT," *Journal of Industrial Information Integration*, vol. 19, no. 2, pp. 100156, 2020.