BEETLE CHIMP OPTIMIZATION ALGORITHM (BCOA) BASED CLUSTER HEAD (CH) ELECTION AND BLOCKCHAIN-BASED DEEP-LEARNING ROUTING IN WIRELESS SENSOR NETWORK (WSN)

Mrs. C. Saranya, Dr. L. Sudha

Research Scholar, Department of Computer Science, VET Institute of Arts and Science College, Erode, TamilNadu, India.

Associate Professor, Department of Computer Science, VET Institute of Arts and Science College, Erode, Tamil Nadu, India.

ABSTRACT: Over the past few years, great importance has been given to Wireless Sensor Network (WSN) as they play a significant role in facilitating the world with daily life services like healthcare, military, social products, etc. However, heterogeneous nature of WSN makes them prone to various attacks, which results in low throughput, and high network delay and high energy consumption. In the WSN, routing is performed using different routing protocols (Low Energy Adaptive Clustering Hierarchy (LEACH), Heterogeneous Gateway-based Energy-Aware Multi-Hop Routing (HMGEAR)). In those protocols, Cluster Head Selection (CHS) plays a vital role in WSN that involves choosing a node to collect and transmit data from a group of sensors. Some nodes in the network may perform malicious activities. In this paper, Energy Efficiency Blockchain Deep Learning (EEBCDL) routing is introduced that accurately classifies all such nodes that depict the same behavior. The WSN contains three types of nodes: sensor nodes, Cluster Head (CH) and the Base Station (BS). Beetle Chimp Optimization Algorithm (BCOA) is introduced for CH selection during clustering process. CHS is constructed based on the residual energy, distance among sensor nodes, CH and BS Distance, Node degree, and Node centrality. The CH after aggregating the data received from the sensor nodes, send it towards the BS. Furthermore, to overcome the single point of failure issue, a decentralized blockchain is deployed on CHs and BS. Additionally, Malicious Nodes (MNs) are removed from the network using Real-Time Message Content Validation (RMCV) and DL techniques like Convolutional Neural Network (CNN), Long Short Term Memory (LSTM), and Gradient Recurrent Unit (GRU) are trained and tested on the LEACH, HMGEAR protocol. Simulation results show that proposed protocol can prolong the network lifetime, improve the network throughput, increased Packet Delivery Ratio (PDR), reduced Packet Loss Ratio (PLR) and reduced average energy consumption. Performance metrics like True Positive Rate (TPR), True Negative Rate (TNR), False Positive Rate (FPR), False Negative Rate (FNR), Overall Accuracy (A) has been used to assess the performance of intrusion detection methods.

INDEX TERMS: Wireless Sensor Network (WSN), Cluster Head Selection (CHS), Energy Efficiency Blockchain Deep Learning (EEBCDL) routing, Beetle Chimp Optimization Algorithm (BCOA), malicious nodes detection, blockchain.

1. INTRODUCTION

The Wireless Sensor Network (WSN) plays a key part in modern era [1]. They acquire data from the surrounding environment and use it for different purposes in different areas like healthcare, smart cities, military, etc. The sensor nodes send the data towards destination without any human involvement in the WSN, after sensing it from the surrounding environment. The data is sent using different communication protocols [2-3]. In general, a node in WSN utilize more energy while data transmission in comparison to processing and sensing. The energy utilization increases manifold when a node start transmitting data to a Base Station (BS) situated at longer distance. The BS is a powerful machine that contains enough storage, processing abilities and also it doesn't have any power constrains [4]. Some approaches have introduced multi-hoping communication to reduce distance covered [5,6]. To enable multi-hoping, several routing protocols are designed that operates on different strategies like data-centric protocols, geographic location-based protocols, clustering and hierarchy-based protocols [7].

The data-centric protocols initiate data transmission between nodes and BS through relay nodes. These types of protocols reduce data redundancy and minimize the amount of data packets transmission. At the same time, these protocols reduce scalability of the network. The geographic protocols communicate data packets to

BS based on sensor node's location. Here the big challenge is to ascertain the geographical location of the nodes. Hierarchical protocols follow multi-tier architecture for data transmission. The whole network is treated like a tree having; normal sensor nodes, Cluster Head (CH) and a BS. Normal nodes collects data and CH aggregate and transmit it to BS. Due to its energy efficiency researchers have focused on designing variety of cluster-based routing protocols to increase network lifetime, load balancing, and scalability.

As an effective scheme to save energy consumption of WSN, a reasonable clustering routing protocol is generally divided into three phases: cluster setup phase, CH election phase, and data transmission phase. In the cluster setup phase, the sensor node groups in the detection area form clusters of different sizes. Based on a certain electoral mechanism, some nodes are selected as the CHs and the remaining nodes act as the member nodes in the CHs election phase. Finally, in the data transmission phase, the member nodes are responsible for collecting environmental information and then transmitting it to the CHs. After the aggregation and data fusion, the CHs send it to the BS. The latter transmits it to the Control Center (CC) via satellite, Internet, or a mobile communication network, eventually the center personnel make decisions based on current environmental information. Figure 1 shows a typical WSN logical hierarchy diagram.

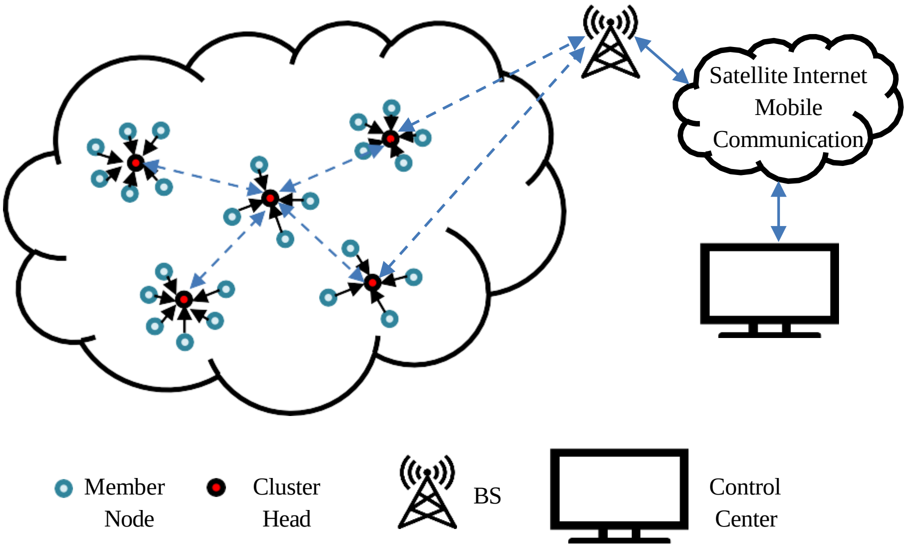

FIGURE 1. CLUSTERING WIRELESS SENSOR NETWORK (CWSN)

In recent years, researchers in various countries have proposed various kinds of clustering protocols for WSN. Becomes more difficult to find that the protocols have been improved in the following aspects: (1) intra-cluster data transmission path, (2) inter-cluster data transmission path, (3) data transmission amount compression [7], (4) implementation of mobile BS or relay node, (5) data security consideration, (6) increase in the number of sink nodes, (7) CHs selection mechanism optimization [8,9]. On the other hand, the problem of unbalanced energy necessitates effective exploitation in the sensor nodes of the clusters for maximizing the partitioning process in order to sustain the challenges that are imposed by the destabilized characteristics of energy in the network. Further, the problem of CH selection in sensor network is a Nondeterministic Polynomial (NP) problem since the energy balanced optimal data aggregation cannot take place in a polynomial time [10-11]. Thus, meta-heuristic optimization algorithms are determined to be potent in facilitating efficient CH selection process [12-13]. Certain contributions have been made with these algorithms and the results have also shown their extraordinary performance in terms of selecting the optimal CH. However, the algorithm often suffers from various drawbacks like poor exploitation phase, slow convergence rate and further it has less capability in solving the multi-objective.

Chimp Optimization Algorithm (ChOA) is a relatively new swarm intelligence-based stochastic technique. It is inspired by the individual intelligence and sexual motivation of chimps in their group hunting [14]. Moreover, the sensor nodes don't generate energy on their own. Therefore, they easily get tampered by the attackers. Blockchain is a promising technology which addresses the data security and third party involvement issues. The blockchain is not only used as a cryptocurrency, but it is also used in various other fields like healthcare, manufacturing, smart cities, etc., [15]. It provides a distributed ledger in which records are added through consensus mechanism after performing the validation process. Moreover, it makes the records more secure without the need of a third party and resolves the issue of single point of failure. In addition, blockchain serves in the field of WSN. The blockchain works as a storage mechanism and provides data security as the sensor nodes are resource constrained. It also works for ensuring nodes' privacy and authentication, non-repudiation as well as performing efficient and secure routing [16], [17].

In this paper, Energy Efficiency Blockchain Deep Learning (EEBCDL) routing is introduced that accurately classifies all such nodes that depict the same behavior. Beetle Chimp Optimization Algorithm (BCOA) is introduced for CH selection during clustering process. CHS is constructed based on the residual energy, distance among sensor nodes, CH and BS Distance, Node degree, and Node centrality. The CH after aggregating the data received from the sensor nodes, send it towards the BS. Furthermore, to overcome the single point of failure issue, a decentralized blockchain is deployed on CHs and BS. Additionally, Malicious Nodes (MNs) are removed from the network using Real-Time Message Content Validation (RMCV) and DL techniques.

2. LITERATURE REVIEW

Zhao et al., [18] introduced a new network structure model and combine the original energy consumption model to construct a new method to determine the optimal number of clusters for the total energy consumption minimization. Based on the balanced energy consumption, then optimize the AGglomerative NESting (AGNES) algorithm, including: (1) introduction of distance variance, (2) the Dual-Cluster Heads (D-CHs) division of the energy balance strategy, and (3) the node dormancy mechanism. In addition, the CHs priority function is constructed based on the residual energy and position of the node. Finally, simulated this protocol in homogeneous networks (the initial energy = 0.4 J, 0.6 J and 0.8 J) and heterogeneous networks (the initial energy = 0.4–0.8 J). Simulation results show that proposed protocol can reduce the network energy consumption decay rate, prolong the network lifetime, and improve the network throughput in the two networks.

Mehta and Saxena [19] presented a Multi-objective Cluster Head Based Energy-aware Optimized Routing (MCH-EOR) algorithm in WSN. Multi-Objective Based Clustering and Sailfish Optimizer (SFO) guided routing method is used to sustain energy efficiency in WSN. CH is selected, based on effective fitness function which is formulated from multiple objectives. It helps to minimize energy consumption and reduces number of dead sensor nodes. After CH selection, SFO is used to select an optimal path to sink node for data transmission. The proposed approach is analytically analyzed and results are compared with the similar existing approaches like Grey Wolf Optimization (GWO), Genetic Algorithm (GA), Ant Lion Optimization (ALO), and Particle Swarm Optimization (PSO) in terms of energy consumption, throughput, PDR, and network lifetime.

Baradaran and Navi [20] proposed a method called the High-Quality Clustering Algorithm (HQCA) for generating high-quality clusters. The HQCA method uses a criterion for measuring the cluster quality, which can improve the inter-cluster and intra-cluster distances as well as reducing the error rate during clustering. The optimal CH is selected based on fuzzy logic and according to various criteria such as the residual energy, the minimum and maximum energy in each cluster, and the minimum and maximum distances between the nodes in each cluster and the base station. The main advantages of this method are its high reliability, low error rate during the clustering process, the independence of key CHs, better scalability, and good performance in large-scale networks with a high number of nodes. The validity of the clustering quality is also measured based on external and internal criteria. Simulation results demonstrated that the HQCA-WSN method can significantly improve the energy consumption and network lifetime. The proposed method also significantly enhances the first node dies and last node dies metrics compared with similar methods.

Khot and Naik [21] proposed an effective and optimal secure routing algorithm named Particle-Water Wave Optimization (P-WWO) for routing the data packets in secure path. The proposed P-WWO algorithm is designed by integrating the Particle Swam Optimization (PSO) with the Water Wave Optimization (WWO). The secure route needed to broadcast the data packets is determined through the selection of Cluster Head using the PSO-based cellular automata with fitness measure. However, the fitness measure is computed by considering the factors, like energy, delay, trust, consistency factor, and maintainability factor. Accordingly, the routing path with the minimal distance and less delay is accepted as the optimal path using the proposed P-WWO based on the fitness value. The route maintenance process enables the proposed optimization to decide whether the packets can be transmitted in the selected route or need to re-route the data.

Adumbabu and Selvakumar [22] presented a dynamic cluster head-based energy-efficient routing system. The Improved Coyote Optimization Algorithm (ICOA) consists of three phases setup, transmission, and measurement phase. The Improved Jaya Optimization Algorithm with Levy Flight (IJO-LF) then determines the route between the BS and the CH. It selects the most effective course based on the distance, node degree, and remaining energy. The proposed approach is compared with traditional methods and the routing protocols Power-Efficient Gathering in Sensor Information Systems (PEGASIS) and Threshold sensitive Energy Efficient Sensor Network protocol (TEEN) during implementation on the MATrix LABoratory (MATLAB) platform. Performance indicators for the suggested methodology are evaluated based on data packets collected by the BS, energy usage, alive nodes, and dead nodes. The outputs of the suggested methodology performed better than the conventional methods.

Zhao et al., [23] proposed a novel Energy-Aware Routing (EAR) approach using Multiobjective Genetic Algorithm and cuckoo (EAR-MOGA-Cuckoo) algorithm for clustering WSN. In this study, clustered nodes are selected from a multiobjective genetic algorithms based on reducing intracluster distances and reducing energy consumption in cluster member nodes and near-optimal routing based on cuckoo optimization algorithm to transfer information between nodes have been used in the direction of the cavity. The implementation results show that considering the evolutionary capabilities of the multiobjective genetic algorithm and the cuckoo optimization algorithm, the proposed method in terms of energy consumption, efficiency, PDR, and packet transmission latency when compared to previous methods.

Mistarihi et al., [24] proposed the performance of several bi-objective optimization algorithms in providing energy-efficient clustering solutions that can extend the lifetime of sensor nodes. Specifically, Moth–Flame Optimization (MFO) algorithm and the Salp Swarm Algorithm (SSA), as well as the Whale Optimization Algorithm (WOA) are considered in providing efficient CH selection decisions. Compared to a reference scheme using the LEACH protocol, the simulation results showed that integrating the MFO, SSA, WOA algorithms into WSN clustering protocols could significantly extend the WSN lifetime, which improved the nodes' residual energy, the number of alive nodes, the fitness function and the network throughput. The results also revealed that the MFO algorithm outperformed the other algorithms in terms of energy efficiency.

Liu et al., [25] proposed a dual-CH clustering routing model and interval-based CH reelection mechanism to reduce the energy consumption associated with frequent cluster formation. In addition, a comprehensive consideration of residual energy and base station distance in CH election facilitates the design of an improved hybrid swarm intelligence optimization algorithm termed GWOA-CH, which combines gray wolf optimization (GWO) and WOA to obtain a more effective CH election scheme. To verify the effectiveness of the proposed algorithm, extensive experiments are conducted between GWOA-CH and some state-of-the-art WSN routing protocols.

Amjad et al., [26] proposed Distance, Degree, and Residual energy-based Low-Energy Adaptive Clustering Hierarchy (DDR-LEACH) protocol. DDR-LEACH is used to replace CHs with the ordinary nodes based on maximum residual energy, degree, and minimum distance from BS. Furthermore, storing a huge amount of data in the blockchain is very costly. Furthermore, for ensuring data security in Interplanetary File System (IPFS), Advanced Encryption Standard (AES) 128-bit is used, which performs better than the existing encryption schemes. Moreover, a huge computational cost is required using a proof of work consensus mechanism to validate transactions. To solve this issue, Proof of Authority (PoA) consensus mechanism is used in the proposed model. The DDR-LEACH is compared with LEACH and the simulation results show that DDR-LEACH outperforms LEACH in terms of energy consumption, throughput, and improvement in network lifetime with CH selection mechanism. Moreover, transaction cost is computed, which is reduced by PoA during data storage on IPFS and service provisioning. Furthermore, the time is calculated in the comparison of AES 128-bit scheme with existing scheme. The formal security analysis is performed to check the effectiveness of smart contract against attacks.

Khan et al., [27] proposed a routing protocol based on Energy Temperature Degree-Low Energy-Adaptive Clustering Hierarchy (ETD-LEACH). In the protocol, nodes consume less energy when transmitting data, which improves the network lifetime. The proposed protocol selects the CHs on the bases of degree, temperature, and energy to perform routing. Blockchain is utilized for solving the issue of a single point of failure. The data transactions are also housed in the blockchain, which is deployed on the CH and BS, as, in blockchain, multiple nodes take part. Therefore, to perform a consensus between them, a PoA consensus mechanism is used in the underlying work. In the blockchain, the Secure Hashing Algorithm-256 (SHA-256) is used for secure hashing of data transactions. Furthermore, malicious nodes are detected during the routing using the Real-time Message Content Validation (RMCV) scheme in the proposed protocol. The proposed model is evaluated under the Denial-of-Service (DoS) attack, the Man-In-The-Middle (MITM) attack, and the smart contract analysis performed by the Oyente tool. ETD-LEACH and Energy THreshold-Low Energy-Adaptive Clustering Hierarchy (ETH-LEACH) protocols are compared using different parameters like number of alive nodes, energy consumption, throughput, and delay. Moreover, PoA transaction cost is less than that of Proof of Work (PoW). Also, the execution time of SHA-256 is less than the execution time of SHA-512.

Yang et al., [28] proposed a trusted routing scheme using blockchain and reinforcement learning to improve the routing security and efficiency for WSN. The feasible routing scheme is given for obtaining routing information of routing nodes on the blockchain, which makes the routing information traceable and impossible to tamper with. The reinforcement learning model is used to help routing nodes dynamically select more trusted and efficient routing links. From the experimental results, it can find that even in the routing environment with 50% malicious nodes, proposed routing scheme still has a good delay performance compared with other routing algorithms. The performance indicators such as energy consumption and throughput also show that trusted routing scheme is feasible and effective.

Tian et al., [29] proposed a Blockchain-based secure Key Management scheme (BC-EKM). First, stake blockchain is constructed based on the hybrid sensor network. In addition, a secure cluster formation algorithm and a secure node movement algorithm is introduced to implement key management, where stake blockchain as a trust machine replaces the majority functions of the BS. BC-EKM scheme is effective and efficient, and better suited to improve the trustworthiness of Dynamic Wireless Sensor Network (DWSN) in the Industrial Internet of Things (IIoT).

3. PROPOSED METHODOLOGY

In this paper, Energy Efficiency Blockchain Deep Learning (EEBCDL) routing is introduced that accurately classifies all such nodes that depict the same behavior. The WSN contains three types of nodes: sensor nodes, Cluster Head (CH) and the Base Station (BS). Beetle Chimp Optimization Algorithm (BCOA) is introduced for CH selection during clustering process. CHS is constructed based on the residual energy, distance among sensor nodes, CH and BS Distance, Node degree, and Node centrality. The CH after aggregating the data received from the sensor nodes, send it towards the BS. Furthermore, to overcome the single point of failure issue, a decentralized blockchain is deployed on CHs and BS. Additionally, Malicious Nodes (MNs) are removed from the network using Real-Time Message Content Validation (RMCV) and DL techniques like Convolutional Neural Network (CNN), Long Short Term Memory (LSTM), and Gradient Recurrent Unit (GRU) are trained and tested on the Low Energy Adaptive Clustering Hierarchy (LEACH) protocol. Simulation results show that proposed protocol can prolong the network lifetime, improve the network throughput, increased Packet Delivery Ratio (PDR), reduced Packet Loss Ratio (PLR) and reduced average energy consumption. Figure 2 shows the overall process of flow diagram.

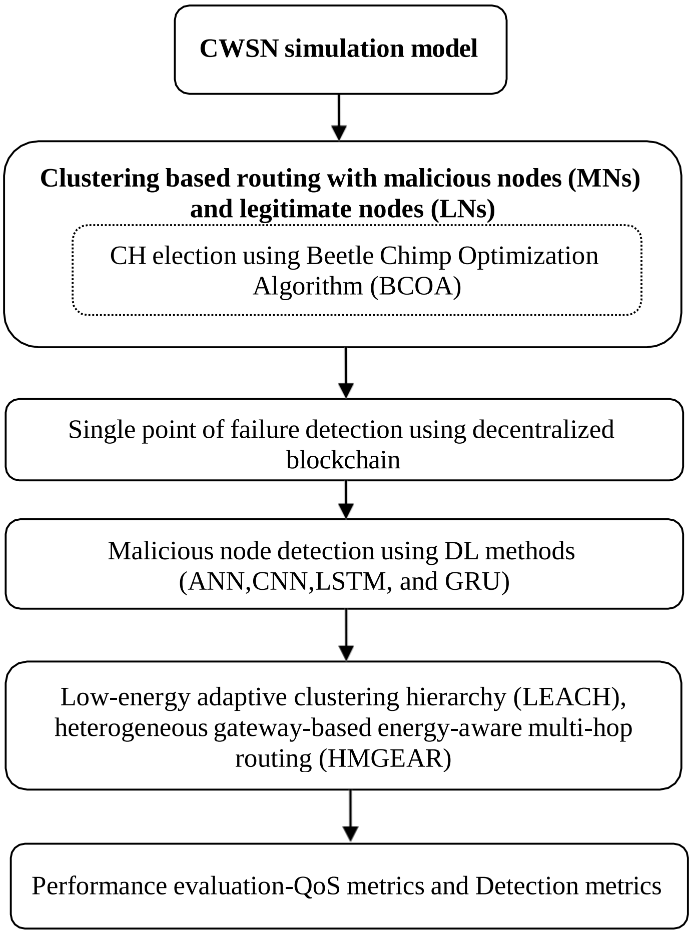
CWSN simulation model
Clustering based routing with malicious nodes (MNs) and legitimate nodes (LNs)
CH election using Beetle Chimp Optimization Algorithm (BCOA)
Single point of failure detection using decentralized blockchain
Malicious node detection using DL methods (ANN,CNN,LSTM, and GRU)
Low-energy adaptive clustering hierarchy (LEACH), heterogeneous gateway-based energy-aware multi-hop routing (HMGEAR)
Performance evaluation-QoS metrics and Detection metrics

FIGURE 2. OVERALL FLOW DIAGRAM OF PROPOSED SYSTEM

3.1. Network Model

Network model, CWSN model is used in which N sensor nodes are evenly arranged in a circular area of diameter M. The BS in the center of the network area has strong computing power. Because the BS energy can be self-replenished, the energy loss of the BS is not considered in this work. It comprising sensor nodes, CHs and BS. On this basis, make the following assumptions about the WSN,

(1) All sensor nodes are static, nodes transmit data to each other in single or multiple hops, and the node energy cannot be supplemented.

(2) The idealized simulation environment does not consider the influence of natural factors such as temperature, humidity, light, and wind on the sensor nodes.

3.2. Energy Consumption Model

This paper quotes the energy consumption model in recent work [23]. According to the actual transmission distance from the CHs to the BS, the free space model and the multipath fading channel model both need to be analyzed by considering only the multipath fading channel model. Therefore, the expression of the total energy consumption of the model will undergo some changes. $E_T(e, d)$ indicates the energy consumed by the wireless transmitter to transmit a set of e bits of information. The expression is as follows:

$$E_T(e, d) = \begin{cases} e \times (E_{elec} + \varepsilon_{fs} d^2), d < d_0 \\ e \times (E_{elec} + \varepsilon_{mp} d^4), d \geq d_0 \end{cases} \tag{1}$$

$E_R(e)$indicates the energy required to receive the information of the e bit. The expression is as follows:

$$E_R(e) = e \times E_{elec} \tag{2}$$

In Equations (1) and (2), $E_{elec}$is the energy consumed per bit by the transmitter or receiving circuit and d is the distance between the transmitter and the receiver. In Equation (2), when $d < d_0$, use the free space model and ε_{fs} acts as the energy factor per bit. Otherwise, the multipath fading channel model is used, and ε_{mp} acts as the energy factor per bit. In addition, $d_0$is used as the distance threshold. As long as it is input as an independent variable into the free space model and the multipath fading channel model to establish an equation (3) the following expression can be obtained:

$$d_0 = \sqrt{\frac{\varepsilon_{fs}}{\varepsilon_{mp}}} \tag{3}$$

In each round of data transmission, the cluster member nodes are responsible for sensing information from the environment, then transmitting it to the CH of the corresponding cluster. Therefore, the calculation formula for the energy consumed by transmitting e bit information is defined as follows:

$$E_{non-CH} = e. E_{elec} + e. \varepsilon_{fs} d^2_{to\,CH} \tag{4}$$

In Equation (4), d_{toCH} represents the distance from the cluster member node to the CH. The CH receives information from the cluster member nodes in the cluster, and then fuses the information that which it senses from the environment, eventually transmits the merged information to the BS. In this paper, let us assume that in each round of data transmission, the information size obtained after processing by the CH is e bit. The energy consumed in the process is calculated as follows:

$$E_{CH} = (\frac{n}{k} - 1)\, e. E_{elec} + \frac{n}{k}\, e. E_{DA} + e. E_{elec} + \begin{cases} e\varepsilon_{fs} d^2_{to\ BS}, d_{to\ BS} < d_0 \\ e\varepsilon_{mp} d^4_{to\ BS}, d_{to\ BS} \geq d_0 \end{cases} \tag{5}$$

the energy consumed consists of three parts: receiving energy consumption, processing energy consumption, and transmitting energy consumption. In Equation (5), n is the number of nodes surviving in the monitored area, k is the number of clusters to be divided, EDA is the energy consumed by the CH to process each bit of data (including received data and sensed data), and $d_{to\ BS}$ is the distance between the CH and the BS.

3.3. Clustering model

In general, the inter-cluster communication traffic in WSN increases as the number of clusters increases, and the intra-cluster communication traffic increases as the number of clusters decreases. In addition, the network's energy consumption increases as communication traffic increases. The determination of the optimal cluster number of the network is of great significance to the network's communication. Determine the optimal number of clusters k in combination with the network structure model and energy consumption model. The main steps of the clustering protocol are described as follows [23]:

(1) Calculate the number k of clusters required according to the calculation formula of the network optimal cluster number.

(2) Through the AGglomerative NESting (AGNES) algorithm with balanced energy consumption optimization, can build the required k clusters.

(3) Implement the selection mechanism of the CH in each cluster, and then can implement the BCOA division of the energy balance strategy and node selection mechanism for the large cluster area before and after the death of the first node, respectively.

(4) Data transmission and energy update.

To minimize the total energy consumption and balance the energy consumption of the nodes in the network, perform the node death decision after each round of data transmission in the network (once the node dies, return to Step 1; otherwise, return to Step 3). In addition, Steps 1, 2, and 3 are collectively called the preparation phase of the protocol. Step 4 is called the stabilization phase of the protocol.

After the sensor network is divided into many clusters, the CH receives the information transmitted by the cluster member nodes, and after processing, eventually transmits it to the BS for final data fusion. In the intra-cluster communication, as the distance between the cluster member node and the CH is not large, adopt the free space model. The distances between some CHs and BS in the model may be larger than d_0, so it is necessary to simultaneously refer to the free space model and the multipath fading channel model when considering the energy consumption between clusters. Before entering the stabilization phase, each member of each cluster needs to send a set of control message named as Nd_{Msg} to its CH in the form of (Nd_{NO}, Nd_{Status}). The status is only divided into work and dormancy. According to the received Nd_{Msg}, the CH of each cluster allocates a time slot for the member nodes of the cluster that need to work. Then every CH sends a set of control message named as ScheduleMsg to its member nodes. Once entering the stabilization phase, the nodes which have received time slots send their sensed information to their associated CHs, and others are in dormancy. As for the CHs, they are responsible for receiving and processing the information sent by the member nodes and eventually transmitting it to the BS. The time slot allocation of the clustering protocol is provided in Figure 3 [23].

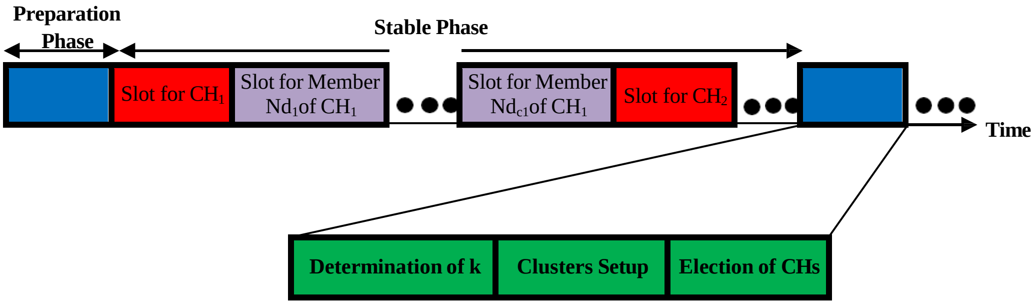

FIGURE 3. TIME SLOT ALLOCATION OF THE CLUSTERING PROTOCOL [23]

The energy consumed by all of the clusters in the region in one round is follows [23]:

$$E_{SUM} = kE_{CH} + nE_{non-CH} \quad (6)$$

A flowchart of the clustering protocol is presented in Figure 4.

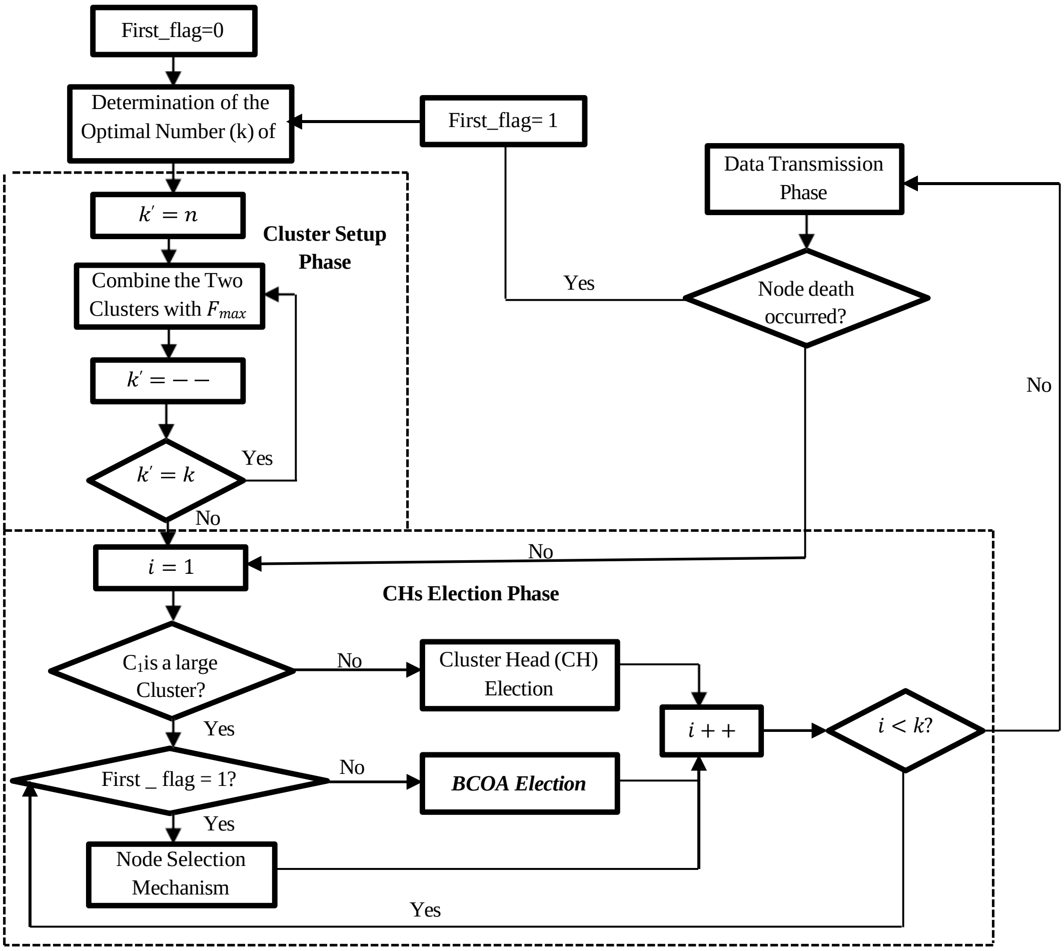

FIGURE 4. FLOWCHART OF THE CLUSTERING PROTOCOL

It includes the determination of k, cluster setup, CH election, and data transmission. The AGglomerative NESting (AGNES) algorithm is a hierarchical clustering algorithm [30]. First, several objects are input, each one constitutes an initial cluster by itself. Then the two clusters with the shortest distance are continuously merged into one cluster until the number of clusters obtained reaches the number of clusters k satisfying the termination condition. Finally, the resulting k clusters are the target clusters of algorithm. In this algorithm, each cluster equals a sample set, and the merger between clusters equals the merger between sets. The merging standard is the distance between the two clusters, which usually assumes three forms: (1) the longest distance, (2) the shortest distance, and (3) the average distance. By adding the two clusters average distance, and the variance of the distance set of two clusters in the cluster setup process, can construct a cluster setup factor. The two clusters corresponding to the largest cluster-merging factor can be merged until the number of clusters reaches the preset number of clusters k [23].

3.3.1. Fitness function

From the sensors group in the network, the optimal CH is selected by the BCOA fitness function. The residual energy is considered in the fitness function to evade a dead node as a CH. Higher centrality is used to reduce the distance of transmission among the members. The mathematical forms of fitness functions and their definitions are discussed below [23]:

CH Residual energy: CH collects data from ordinary sensor nodes and transmits it to BS in a network. Because the CH takes more energy to perform the preceding activities, the node with the most residual energy is the strongest choice to be a CH. The following Equation (7) describes residual energy (f_1) [23]:

$$f_1 = \sum_{i=1}^{m} \frac{1}{E_{CH_i}} \tag{7}$$

where the ith CH residual energy is E_{CH_i}.

Distance among sensor nodes: It specifies the range among usual sensor nodes as well as its CH. The dissipation of energy for the node mostly depends on the transmission path distance. If the chosen node has a minimum transmitting distance near BS, then the node’s energy consumption is small. Sensors to CH (f_2) the distance is displayed in Equation (8) [23]:

$$f_2 = \sum_{j=1}^{m} \sum_{i=1}^{I_j} D\left(s_i, \frac{CH_j}{I_j}\right) \tag{8}$$

where $D\left(s_i, \frac{CH_j}{I_j}\right)$ denoted the sensor i and CH_j distance and I_j is the sensor node’s quantity belongs to CH [23].

CH and BS Distance: The energy consumption of the node is evaluated based on the distance via the transmitting track. When BS is situated away from CH, for example, data transmission will require a lot of energy. As a result, the abrupt drop in CH could be related to increased energy use. Therefore, throughout information transfer, the node that is closest to BS is selected. The optimal solution of distance among the BS (f_3) and the CH is represented by the following Equation (9) [23]:

$$f_3 = \sum_{i=1}^{m} D(CH_j, BS) \tag{9}$$

where $D(CH_j, BS)$is the distance between BS and CH_j.

Node degree: It specifies how many sensor nodes each CH has. Because CHs with more cluster members lose their energy for a shorter period, the CHs with fewer sensors are chosen. Equation (10) expresses the degree of node (f_4) [23]:

$$f_4 = \sum_{i=1}^{m} I_i \tag{10}$$

where $I_i$is the number of CHi sensor nodes.

Node centrality: Node centrality (f_5) is an expression that expresses how far a node is from its neighbors, represented in Equation (11) [23]:

$$f_5 = \sum_{i=1}^{m} \frac{\sqrt{\frac{\left(\sum_{j\in n}^{m} D^2(i,j)\right)}{n(i)}}}{L} \tag{11}$$

where n(i) is the number of CH_i neighboring nodes and L is the network dimension. Every objective value is assigned a weight value. Several objectives are combined into a single function in this situation. S1, S2, S3, S4, and S5 are the weighted values. Equation (12) depicts the single objective function [23]:

$$F = S1f_1 + S2f_2 + S3f_3 + S4f_4 + S5f_5 \tag{12}$$

where, $\sum_{r=1}^{5} Sr = 1\, Sr \in (0,1)$, the values of Sr are 0.30, 0.25, 0.2, 0.15 and 0.1 correspondingly [23].

3.3.2. CH Selection Using BCOA

Chimp Optimization Algorithm (COA) is a metaheuristic algorithm which simulates the social status relationship and hunting behavior of chimps [30-31]. BCOA is proposed to improve the accuracy of solutions. First, the disruptive polynomial mutation is used to initialize the population, which provides the foundation for global search. Next, reduce the probability of falling into the local optimum, the beetle antennae operator is used to improve the less fit chimps while gaining visual capability.

Its intuitive background originates from the hunting behavior of chimps. Chimps perform different actions according to the division of labor to find the prey. COA algorithm divides the chimp group into four types: attacker, barrier, chaser, and driver. Among them, the attacker is the leader of the population. The other three types of chimps assisted in hunting, and their status decreased in turn.

Sensor nodes parameter and random initialization: Initially, the node’s global population, including Np packets, each with Ns nodes are initialized. The mathematical model is briefly described as following. Equations (13) and (14) is used to update the position of the chimp [30].

$$\begin{gathered} X_1(t+1) = X_{Attacker}(t) - a_1 . d_{Attacker} \\ X_2(t+1) = X_{Barrier}(t) - a_2 . d_{Barrier} \\ X_3(t+1) = X_{Chaser}(t) - a_3 . d_{Chaser} \\ X_4(t+1) = X_{Driver}(t) - a_4 . d_{Driver} \end{gathered} \tag{13}$$

$$X_{Chimp}(t+1) = \frac{X_1 + X_2 + X_3 + X_4}{4} \tag{14}$$

where t represents the number of the current iteration, the position of the chimp is updated according to the four types of position depending on the packets and sensor nodes. The dynamic coefficient a and vector d are expressed in equation (15):

$$\begin{aligned} a_1 &= 2f_1r_1 - f_1, d_{Attacker} = |c.X_{Attacker}(t) - m.X(t)| \\ a_2 &= 2f_2r_1 - f_2, d_{Barrier} = |c.X_{Barrier}(t) - m.X(t)| \\ a_3 &= 2f_3r_1 - f_3, d_{Chaser} = |c.X_{Chaser}(t) - m.X(t)| \\ a_4 &= 2f_4r_1 - f_4, d_{Driver} = |c.X_{Driver}(t) - m.X(t)| \end{aligned} \tag{15}$$

where the coefficient f decreases nonlinearly from 2.5 to 0 with the lapse of iteration. $c = 2r_2.r_1$ and $r_2$are random numbers in [0, 1]. m is a chaotic map vector. Assuming the probability μ is a random number in [0, 1], the chaotic model is used for position updating when $\mu \geq 0.5$, as shown in equation (16). Otherwise, equation (14) is still executed:

$$X_{chimp}(t+1) = \text{Chaotic\_Value} \tag{16}$$

Common mutation operators include random mutation, nonuniform mutation, and polynomial mutation. For the traditional polynomial mutation (PM), the mutation has no effect when the variable is on the boundary. Highly Disruptive Polynomial Mutation (HDPM) improves this disadvantage [32]. The form of the operator is in equation (17):

$$X_{new} = X + \delta_k.(\text{ub-lb}) \tag{17}$$

where ub and lb represent the upper and lower boundaries of the search space depending on the packets and sensor nodes with distance, residual energy, degree and centrality. X is the parent, $X_{new}$is the offspring. The coefficient δ_k is calculated by Equations (18)-(20):

$$\delta_1 = \frac{X - lb}{ub - lb} \tag{18}$$

$$\delta_2 = \frac{ub - X}{ub - lb} \tag{19}$$

$$\delta_k = \{ \begin{array}{ll} [2r + (1-2r).(1-\delta_1)^{\eta_m+1}]^{\frac{1}{\eta_m+1}} - 1, & \text{if } r \leq 0.5 \\ 1 - [2(1-2r) + 2(r-0.5).(1-\delta_2)^{\eta_m+1}]^{\frac{1}{\eta_m+1}}, & \text{else} \end{array} \tag{20}$$

where r is a random number in [0, 1], $\eta_m$is the mutation index. It can be observed from Equation (20) that it can still make full use of the whole search space of CH. This advantage maintains the diversity of candidate solutions as CH. When a beetle is preying, it receives the food smell of near-area using two antennas and finds the area with the strongest smell (distance). If the antennae on one side receive a stronger concentration of smell, the beetle will turn to the same side (CH); otherwise, it will turn to the other side (CH). Considering the sense of smell, the beetle antennae search algorithm [33-34]. First, the random direction vector $\vec{b}\rightarrow$ is normalized to the following form:

$$\vec{b}\rightarrow = \frac{rnd(n,1)}{\|rnd(n,1)\|} \tag{21}$$

where rnd(.) represents a random function, and n represents the dimension of CH search space. The two search behaviors in Equation s(22-23) simulates the beetle exploring the left and right areas using two antennas:

$$X_r(t) = X(t) + d(t).\vec{b}\rightarrow \tag{22}$$

$$X_l(t) = X(t) - d(t).\vec{b}\rightarrow \tag{23}$$

where *X* represents the original node position of the beetle, $X_r$represents the best CH position obtained by the beetle exploring the right areas, $X_l$represents the CH position obtained by the beetle searching the left areas, *d* represents the distance between two antennas. The update rule is as follows:

$$d(t) = \frac{\delta(t)}{C} \tag{24}$$

where δ represents the step size, *C* represents the attenuation rate, *C* = 2. δ is calculated in Equation (25):

$$\delta(t) = K.\delta(t-1) \tag{25}$$

where K is the attenuation rate of δ, K=0.95 [34]. Based on the above search behaviors, the new position update model of the beetle is shown in Equation (26):

$$\mathrm{X}(\mathrm{t}+1)=\mathrm{X}(\mathrm{t})+\delta(\mathrm{t}).\vec{b}\rightarrow.sign\ (f(X_r(t)-f(X_l(t))) \qquad (26)$$

where sign(.) represents a sign function. The pseudo-code is given in Algorithm 1,

Algorithm 1. Pseudocode of BCOA

```
Initialize the population size N with number of sensor nodes Ns and the maximum number of iterations $Max_{iter}$
Initializes positions of chimps $X$ including Np packets, each with Ns nodes
Generate mutant positions of chimps using equation (17)// HDPM//
Calculate the fitness of each chimp using equation (12)
Select attacker, barrier, chaser, and driver
While $t < Max_{iter}$
    For $i$=driver to $N$
        For$j$=1 to $n$
            Calculate fitness of each chimp using equation (12)
            If $p \leq 0$
                Update the position of $X$ using equation (26) //Introduce beetle antennae operator//
            End if
        End for
    End for
    For $i = 1$ to $n$
        If μ<0.5
            Update the CH position of the current chimp using equation (14)
        Else if
            Update the CH position of the current chimp using equation (15)
        End if
    End for
    Update $f, c, m, a$ and $d$
    Calculate the fitness of all chimp using equation (12)
    Update $X_{Anacker}, X_{Barrier}, X_{Chaser}, and\ X_{Driver}$
    t=t+1
End while
Return $X$
```

First, the HDPM strategy is introduced to enhance the population diversity in the initialization phase of node and their packets. It should achieve an appropriate balance between exploration and exploitation. Further, high exploration steps are carried out at the beginning of the search of CH, and more exploitation steps are often required in the last phase. Second, Beetle antennae operator, the less fit chimp gains visual capability. The node position updating method of chimp is improved using Equation (26). It utilizes its left and right eyes to observe the distance on both sides and further determines the direction of the next step. The purpose of this improvement is to prevent the driver chimp from falling in the local optimum due to poor performance. The flowchart is shown in Figure 5.

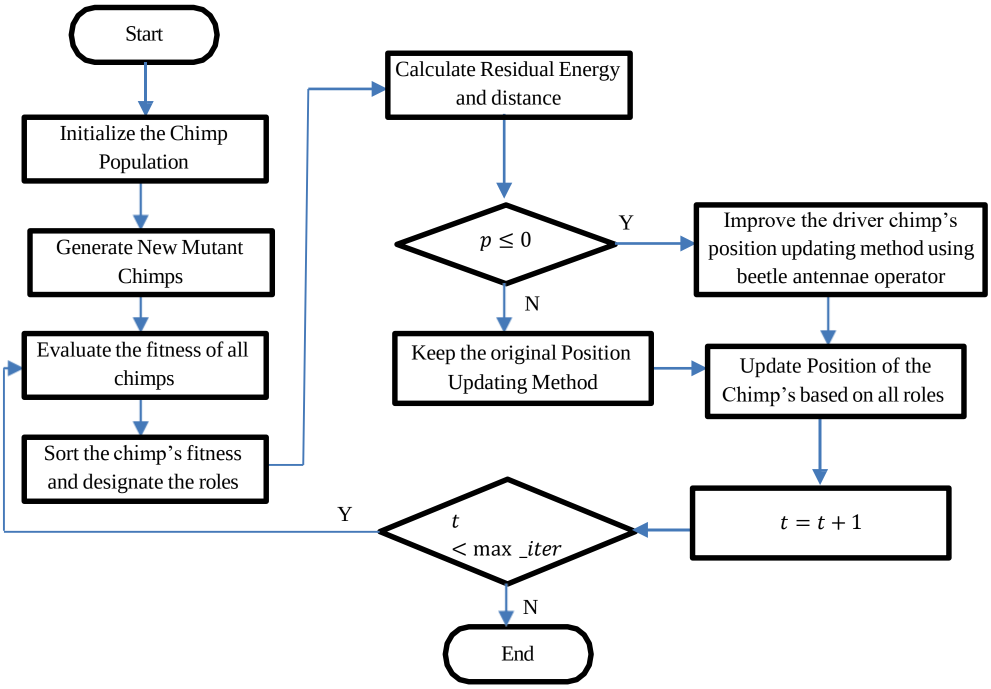
Start
Initialize the Chimp Population
Generate New Mutant Chimps
Evaluate the fitness of all chimps
Sort the chimp's fitness and designate the roles
Calculate Residual Energy and distance
$p \leq 0$
Y
N
Improve the driver chimp's position updating method using beetle antennae operator
Keep the original Position Updating Method
Update Position of the Chimp's based on all roles
$t = t + 1$
$t < \max\_iter$
Y
N
End

FIGURE 5. FLOWCHART OF BCOA

3.4. THREE-LAYERED ARCHITECTURE

This section presents a three-layered architecture, as shown in Figure 6. The middle layer is the blockchain enabled RMCV and DL network layer. Whereas, the first and the third are the CWSN layers, comprising sensor nodes, CHs and BS. The BS and CHs are the nodes where blockchain is deployed. Moreover, ANN, CNN, long short term memory (LSTM), and gradient recurrent unit (GRU) are trained and tested on the LEACH protocol generated dataset. The nodes in the CWSN perform routing using both LEACH and HMGEAR protocols. During routing, some nodes in the network may perform malicious activities. To minimize such activities in the network, the MNs are detected using the trained model and RMCV scheme. The nodes involved in the routing mechanism are classified using DL and RMCV techniques. Once classification is performed, the registration of LNs is done using PoA. While, routing is performed using LEACH and HMGEAR protocols after MND.

BLOCKCHAIN NETWORK: The blockchain enabled RMCV and CNN detect the presence of MNs in the WSN. It is deployed on CHs and BS. After MND, the LNs are registered in the blockchain using the smart contract. The PoA consensus mechanism is used when LNs are registered in blockchain network.

CWSN: The CWSN is composed of N number of nodes that perform routing using LEACH and HMGEAR protocols. After data processing, the data is collected by the sensor nodes from the surrounding environment and sent towards the CHs for processing. Further data processing is performed by the CHs and the integrated data is sent towards BS. In the WSNs, some nodes may perform malicious activities and may send false messages to disturb the data traffic or drop the data packets. Therefore, MND is performed using RMCV and DL-driven architecture. RMCV calculates the nodes' trustworthiness; other side nodes' classification is performed by ANN, CNN, LSTM, and GRU. The RMCV scheme detects the MNs in the WSNs. In the WSNs, the victim nodes may find the adversary messages sent by adversary nodes. The messages are evaluated based on their trustworthiness and integrity parameters, which are calculated by satisfying the following three conditions: message path, message conflict and message similarity. Each message contains a trust value and based on that trust value, the trust of each node is calculated. The nodes that send the maximum number of trusted messages are considered legitimate and trusted. While the nodes that send the minimum number of trusted messages are considered malicious and untrusted. The MNs are removed from the network after being detected. While, routing is performed only between the LNs using LEACH and HMGEAR protocols.

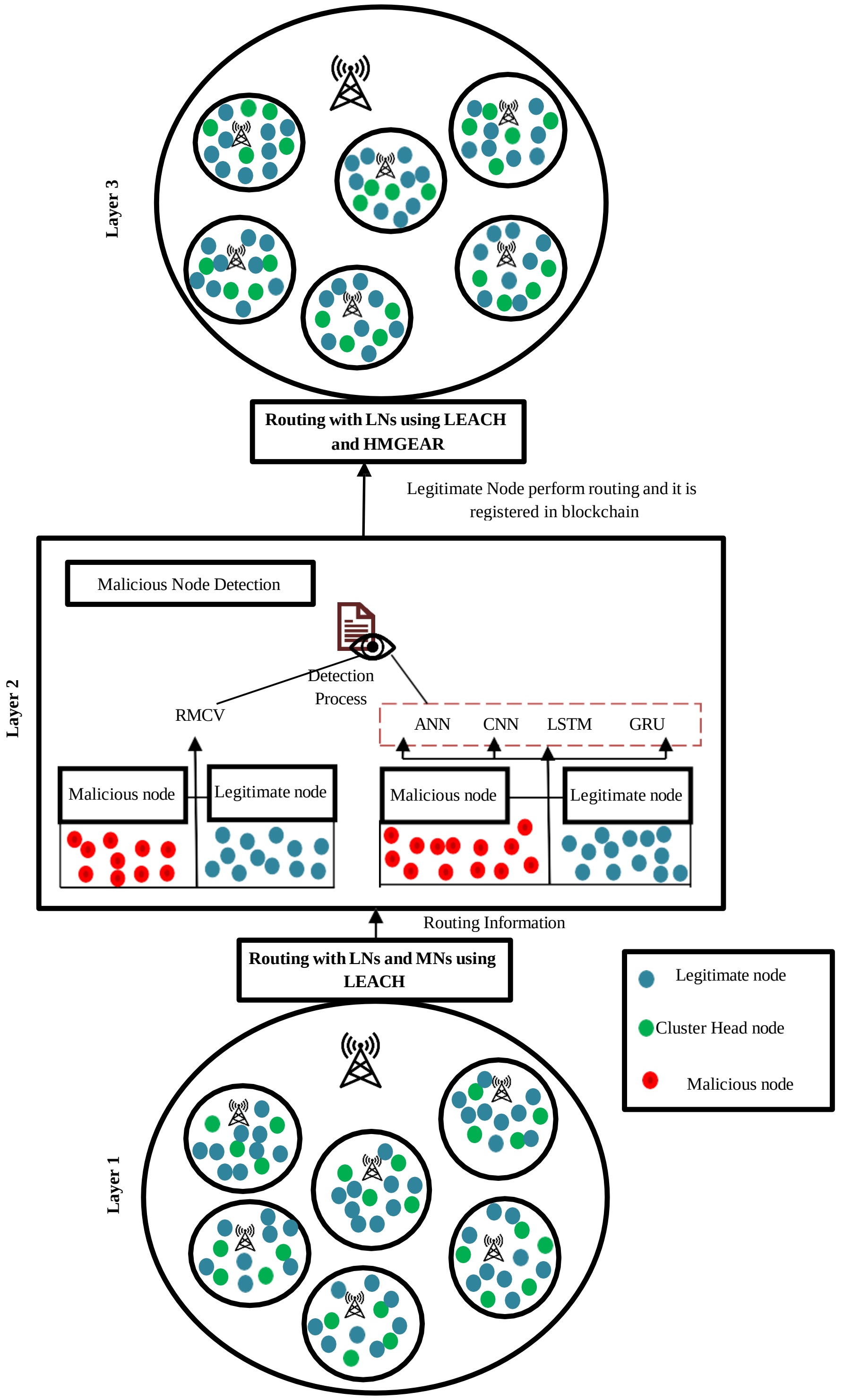

FIGURE 6. NEWLY PROPOSED 3-LAYERED SYSTEM MODEL

While the ANN contains only three layers: input, hidden, and output. The hidden layer processes the given input and forwards it to the fully connected neurons where the output is generated [40]. Therefore, CNN gives more accuracy than ANN. CNN is a regularized type of Feed-Forward Neural Network (FFNN) that learns features by itself via filter (or kernel) optimization. FFNN are usually fully connected networks, that is, each neuron in one layer is connected to all neurons in the next layer. The "full connectivity" of these networks makes them prone to overfitting data. Typical ways of regularization, or preventing overfitting, include: penalizing parameters during training (such as weight decay) or trimming connectivity (skipped connections, dropout, etc.) Robust datasets also increase the probability that CNNs will learn the generalized principles that characterize a given dataset rather than the biases of a poorly-populated set. It is a multi layered architecture where the layers work well for feature extraction and classification purposes [37-38].

In the proposed work, LSTM is employed for classification purpose. The issue of gradient vanishing using explicit memory unit is also solved by LSTM. Moreover, anything can be memorized in LSTM using any weights. In this way, LSTM memorizes and performs better than the previous models. LSTM, the gated mechanism is used to store and pass the information to the next layer. Moreover, LSTM has another feature for smooth and uninterrupted flow of gradients during propagation, which is called constant error carousel [39-40]. GRU works like LSTM as it uses gates for the information control. GRU is basically an updated version of LSTM having some improvements [41]. The information is transferred towards output using two vectors. These vectors decide which information should be passed to the output. To keep the information for a long time, these vectors are trained. However, GRU doesn't remove irrelevant information. As compared to LSTM, GRU has only one hidden cell state. Therefore, it is quicker to train as compared to LSTM [42].

GRU solves gradient vanishing problem using update and reset gate. If the gradient shrinks, then it back propagates over time and affects the learning, which makes the model untrainable. GRU gives the advantage over LSTM as it uses less memory and works better than LSTM. GRU also works faster than LSTM. On the other side, the LSTM works better on the datasets having long sequences. DL models find the maximum MNs in the network and remove them. Moreover, the Legitimate Nodes (LNs) identified by DL models are registered in the blockchain network using a smart contract.

In the proposed model, the time for transmission, in the tth communication round, to the BS from the nth client is given using Equation (27) [43]:

$$\tau_n^{up}(t) = \frac{\gamma_n i_n(t)}{B\log_2\left(1 + \frac{P_n(t)h_n(t)}{BN_0}\right)} \tag{27}$$

where the system bandwidth, transmission power of the nth node, noise power spectral density and the number of bits are given by nth node are given by B, $P_n(t)$, N_0 and γ_n. In addition, Equation (28) gives the energy consumed by the nth node in the tth communication round [43].

$$E_n^{up}(t) = \frac{P_n(t)\gamma_n i_n(t)}{B\log_2\left(1 + \frac{P_n(t)h_n(t)}{BN_0}\right)} \tag{28}$$

4. SIMULATION RESULTS AND DISCUSSION

Simulation WSN-DS, a specialized dataset for WSN to detect DoS attacks. LEACH and protocol was used to collect the dataset because it is one of the most common and widely used routing protocols in WSNs. WSNDS contains 374661 records that represent four types of DoS attacks: Blackhole, Grayhole, Flooding, and Scheduling attack, in addition to the normal behavior (no-attack) records (Table 1). The dataset is separated to 60.00% training data and 40.00% testing data. Table 1 shows data separation using holdout method. It has been collected from https://www.kaggle.com/datasets/bassamkasasbeh1/wsnds.Network Simulator 2 (NS2) was used to gather the required data [45] and simulate the model. Simulation parameters are summarized in Table 2. Experiments are conducted using hardware platform of Intel (R) Xeon (R), i5-3570 CPU with the clock frequency of 3.50 GHz and 16.00 GB memory, 10 MB cache. Simulation is performed for different number of sensor nodes various percentage of CHs at different locations of BS in the sensing field of area $200 \times 200\ m^2$. The sensor nodes change within a range of 100 to 400 and the number of selected CHs vary from 15% to 30% while BS is positioned at (50,50), (100, 100), (150, 150).

TABLE 1. WSN-DS DATASET DETAILS

ATTACK	TRAINING SET	TESTING SET
Blackhole	6029	4020
Grayhole	8758	5838
Flooding	1988	1324
Scheduling	3982	2656
Normal	204039	136027
Total	224796	149865

TABLE 2.VARIOUS PARAMETERS USED IN SIMULATION

Parameter	Value
Network area	$200 \times 200\ m^2$
Base Station location	(50,50), (100,100), (150,150)
Number of sensor nodes	500 nodes

Number of clusters	20
Initial Sensor Node Energy	2.0 J
$\mathbf{E_{elec}}$	50 nJ/bit
$\boldsymbol{\varepsilon_{fs}}$	10 pJ/bit/m^2
$\boldsymbol{\varepsilon_{mp}}$	0.0013 pJ/bit/m^4
$\boldsymbol{d_0}$	30 m
$\boldsymbol{d_{max}}$	200 m
Packet Size	4000 bits
% of CHs	10% to 30%
Routing Protocol	LEACH, HMGEAR
MAC Protocol	CSMA/TDMA
Attacker intensity	10%, 30%, 50%
Network Topology	Random distribution
Number of rounds	3000

For better CH selection, distance to BS, node centrality, residual energy, node degree, and distance to neighbors are the inputs given to BCOA. Distance, node degree, and residual energy are also inputs to the EEBCDL. Because these methods are commonly employed to enhance the WSNs energy efficiency, this proposed method (BCOA-EEBCDL) is contrasted to several established approaches such as the DDR-LEACH [26], ETD-LEACH [27], MCH-EOR [19], EAR-MOGA-Cuckoo [23].

QoS Metrics: Simulation results show that proposed protocol can prolong the network lifetime, improve the network throughput, increased Packet Delivery Ratio (PDR), reduced Packet Loss Ratio (PLR) and reduced average energy consumption.

Throughput: Throughput is the amount of data that passes through the network in a given time period. It's usually measured in bits per second (bps) or megabits per second (Mbps).

Packet Delivery Ratio (PDR): Packet Delivery Ratio (PDR) can be expressed as follows,

$$R_{Success} = \frac{P_{succ}}{P_{all}} \times 100\% \tag{29}$$

where $R_{Success}$ represents the ratio of packet successful, P_{succ} denotes the number of packets successfully sent by the router, P_{all} indicates the total number of packets transmitted through the router.

Packet Loss Ratio (PLR): Packet Loss Ratio (PLR) can be expressed as follows,

$$R_{drop} = \frac{P_{drop}}{P_{all}} \times 100\% \tag{30}$$

where R_{drop}represents the ratio of packet loss, P_{drop} denotes the number of packets dropped by the router, P_{all} indicates the total number of packets transmitted through the router.

Average energy consumption: During each iteration, it specifies how much energy each node uses on average. It has been measured in terms of mJ.

Detection metrics: Because different performance metrics are appropriate in different settings, performance metrics like True Positive Rate (TPR), True Negative Rate (TNR), False Positive Rate (FPR), False Negative Rate (FNR), Overall Accuracy (A) has been used to assess the performance of intrusion detection methods. TPR represents the rate of attack cases identified correctly, TNR represents the rate of normal (no-attack) cases identified correctly, FPR represents the rate of no-attack cases identified as attacks by the system, and FNR represents the rate of attack cases identified as normal ones. A is the total rate of correct decisions whether identifying an attack correctly or deciding there is no attack when really there is no attack.

$$\text{True Positive Rate (TPR)} = \frac{TP}{TP + FN} \times 100\% \tag{31}$$

$$\text{True Negative Rate (TNR)} = \frac{TN}{TN + FP} \times 100\% \tag{32}$$

$$\text{False Positive Rate (FPR)} = \frac{FP}{FP + TN} \times 100\% \tag{33}$$

$$\text{False Negative Rate (FNR)} = \frac{FN}{FN + TP} \times 100\% \tag{34}$$

$$\text{A} = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \tag{35}$$

where TP is the number of attack cases classified correctly as attacks. TN is the number of normal (no-attack) cases classified correctly as normal (no-attack). FP is the number of normal (no-attack) cases classified incorrectly as attacks. FN is the number of attack cases classified incorrectly as normal (no-attack). Attack detection results of DL methods with respect to evaluation metrics are illustrated in table 3.

TABLE 3. PERFORMANCE METRICS OF INTRUSION DETECTION SYSTEMS

METHODS	TPR (%)	TNR (%)	FPR (%)	FNR (%)	A (%)
ANN	84.12	85.61	14.39	15.88	85.37
CNN	86.69	87.45	12.55	13.31	87.62
LSTM	89.54	90.89	9.11	10.46	90.41
GRU	91.49	92.08	7.92	8.51	92.17

TABLE 4. PERFORMANCE ANALYSIS COMPARISON OF ROUTING METHODS

Number of nodes	**Network lifetime (rounds)**				
	DDR-LEACH	**ETD-LEACH**	**MCH-EOR**	**EAR-MOGA-Cuckoo**	**BCOA-EEBCDL**
100	892	1058	1207	1412	1645
200	1241	1395	1571	1732	1921
300	1625	1753	1922	2047	2258
400	1858	2026	2148	2334	2587
500	2215	2377	2519	2693	2814
Number of nodes	**Throughput(Mbps*10^4)**				
	DDR-LEACH	**ETD-LEACH**	**MCH-EOR**	**EAR-MOGA-Cuckoo**	**BCOA-EEBCDL**
100	8.22	10.18	11.85	13.44	15.73
200	10.85	12.97	14.75	16.61	18.47
300	13.71	15.46	17.38	18.76	20.32
400	15.68	17.05	18.76	20.14	21.83
500	17.45	18.63	20.11	21.62	23.75
Number of nodes	**Packet Delivery Ratio(PDR)(%)**				
	DDR-LEACH	**ETD-LEACH**	**MCH-EOR**	**EAR-MOGA-Cuckoo**	**BCOA-EEBCDL**
100	82.18	84.59	86.63	88.75	90.52
200	80.66	83.21	85.46	87.92	89.71
300	79.72	81.47	84.18	86.79	88.54
400	78.26	80.43	83.09	85.61	87.46
500	77.35	79.24	81.26	84.59	86.47
Number of nodes	**Packet Loss Ratio(PLR)(%)**				
	DDR-LEACH	**ETD-LEACH**	**MCH-EOR**	**EAR-MOGA-Cuckoo**	**BCOA-EEBCDL**
100	17.82	15.41	13.37	11.25	9.48
200	19.34	16.79	14.54	12.08	10.29
300	20.28	18.53	15.82	13.21	11.46
400	21.74	19.57	16.91	14.39	12.54
500	22.65	20.76	18.74	15.41	13.53
Number of Rounds	**Average Energy Consumption (mJ)**				
	DDR-LEACH	**ETD-LEACH**	**MCH-EOR**	**EAR-MOGA-Cuckoo**	**BCOA-EEBCDL**
500	0.362	0.335	0.301	0.272	0.234
1000	0.394	0.359	0.327	0.298	0.262
1500	0.425	0.397	0.364	0.327	0.284
2000	0.458	0.423	0.387	0.351	0.298
2500	0.487	0.456	0.415	0.372	0.321
3000	0.529	0.504	0.468	0.429	0.358

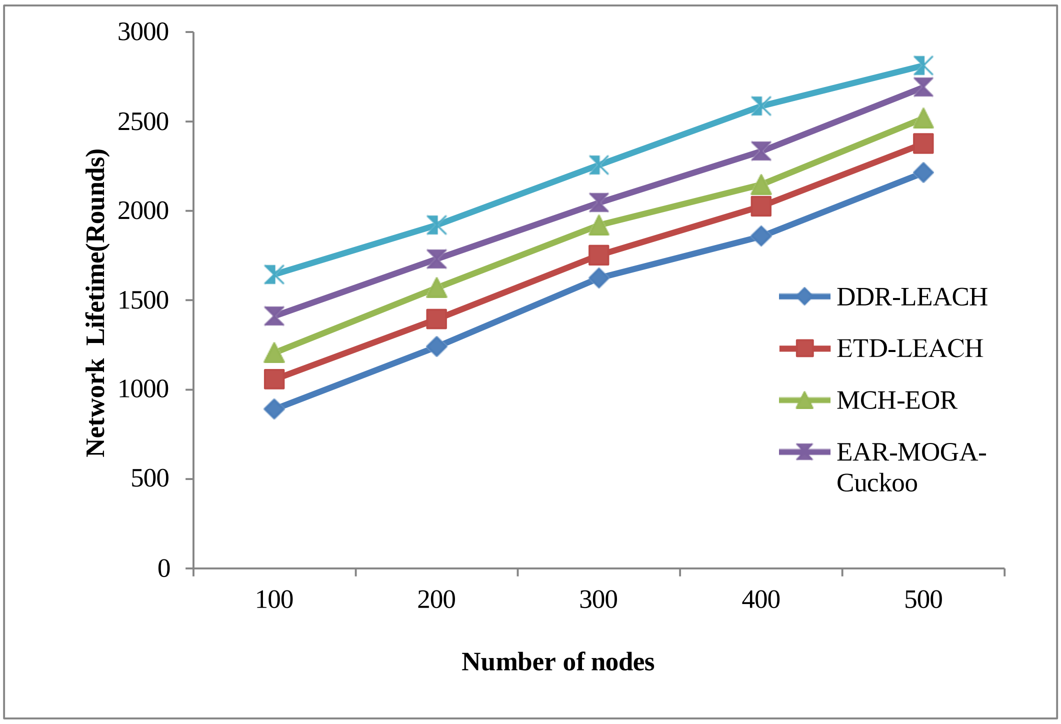

FIGURE 7. NETWORK LIFETIME VS. NUMBER OF NODES

Figures 7 illustrate the proposed method network lifetime in terms of rounds with conventional methods like DDR-LEACH, ETD-LEACH, MCH-EOR, and EAR-MOGA-Cuckoo. The graph shows a good improvement in the network lifetime of the proposed method compared with existing methods of network size 100 to 500 nodes. It shows that the proposed system has highest network lifetime of 2814 rounds, other methods such as DDR-LEACH, ETD-LEACH, MCH-EOR, and EAR-MOGA-Cuckoo has lowest network lifetime of 2215 rounds, 2377 rounds, 2519 rounds, and 2693 rounds for 500 nodes (See Table 4).

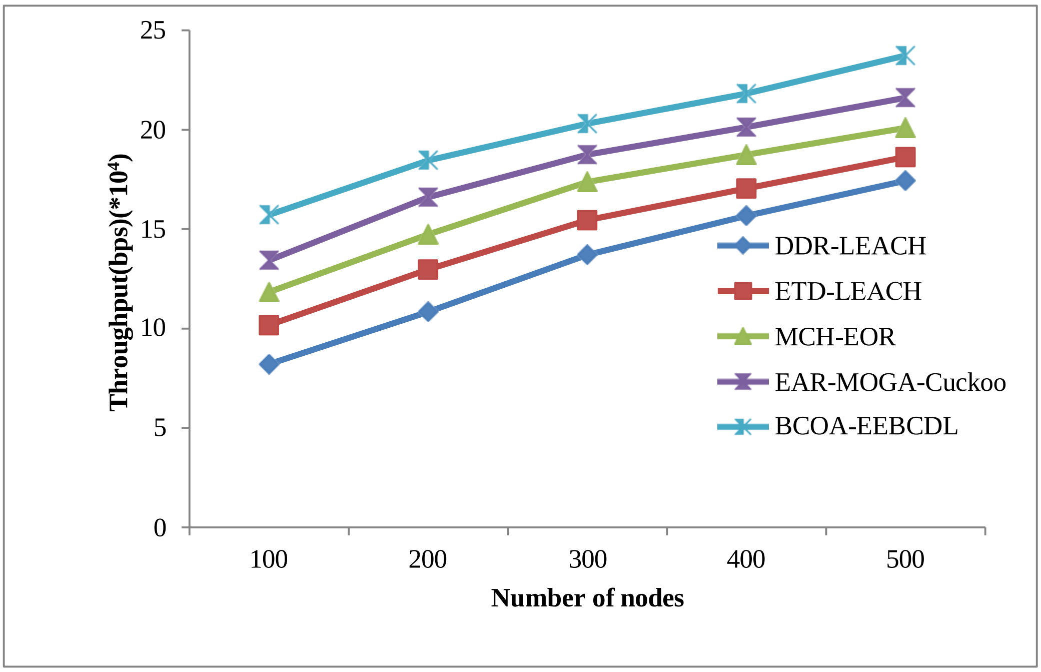

FIGURE 8. THROUGHPUT COMPARISON VS. ROUTING PROTOCOLS

Proposed method and conventional methods like DDR-LEACH, ETD-LEACH, MCH-EOR, and EAR-MOGA-Cuckoo with respect to throughput is illustrated in figure 8. The graphs show a greater improvement in throughput with the increase in sensor nodes. It shows that the proposed system has highest throughput of 23.75*10^4 bps, other methods such as DDR-LEACH, ETD-LEACH, MCH-EOR, and EAR-MOGA-Cuckoo has lowest throughput of 17.45*10^4 bps, 18.63*10^4 bps, 20.11*10^4 bps, and 21.62*10^4 bps for 500 nodes (See Table 4).

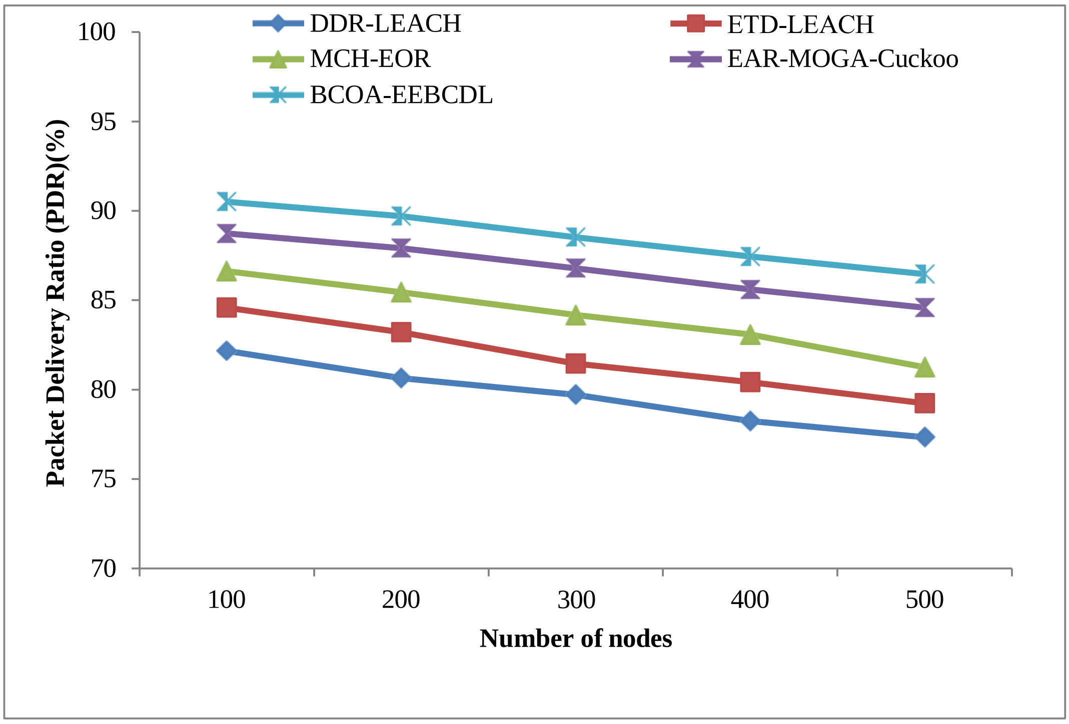

FIGURE 9. PDR COMPARISON VS. ROUTING PROTOCOLS

PDR results of DDR-LEACH, ETD-LEACH, MCH-EOR, EAR-MOGA-Cuckoo and BCOA-EEBCDL with 100 and 500 nodes are shown in Figure 9. Proposed system has highest PDR of 90.52%, other methods such as DDR-LEACH, ETD-LEACH, MCH-EOR, and EAR-MOGA-Cuckoo has lowest PDR of 82.18%, 84.59%, 86.63%, and 88.75% for 100 nodes (See Table 4).

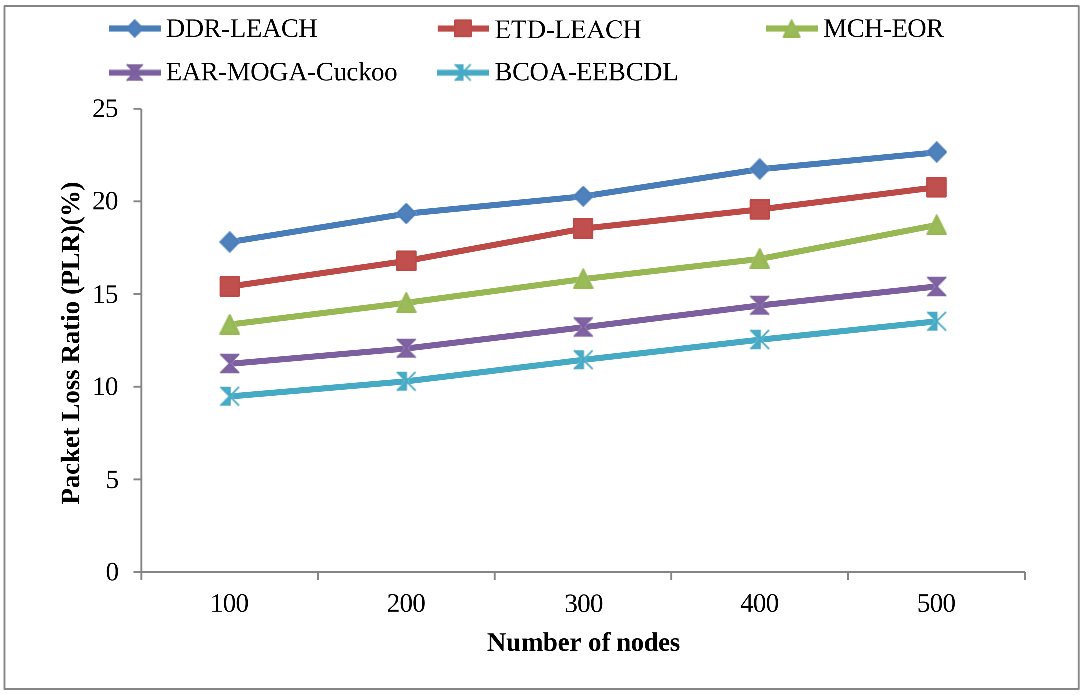

FIGURE 10. PLR COMPARISON VS. ROUTING PROTOCOLS

Figure 10 depict the comparison of the BCOA-EEBCDL with existing techniques like DDR-LEACH, ETD-LEACH, MCH-EOR, and EAR-MOGA-Cuckoo in terms of packet drop ratio with 100 to 500 nodes. PLR of the proposed method is compared with existing methods of network size 100 to 500 nodes. Proposed system has lowest PLR of 9.48%, other methods such as DDR-LEACH, ETD-LEACH, MCH-EOR, and EAR-MOGA-Cuckoo has increased PLR of 17.81%, 15.41%, 13.37%, and 11.25% for 100 nodes (See Table 4).

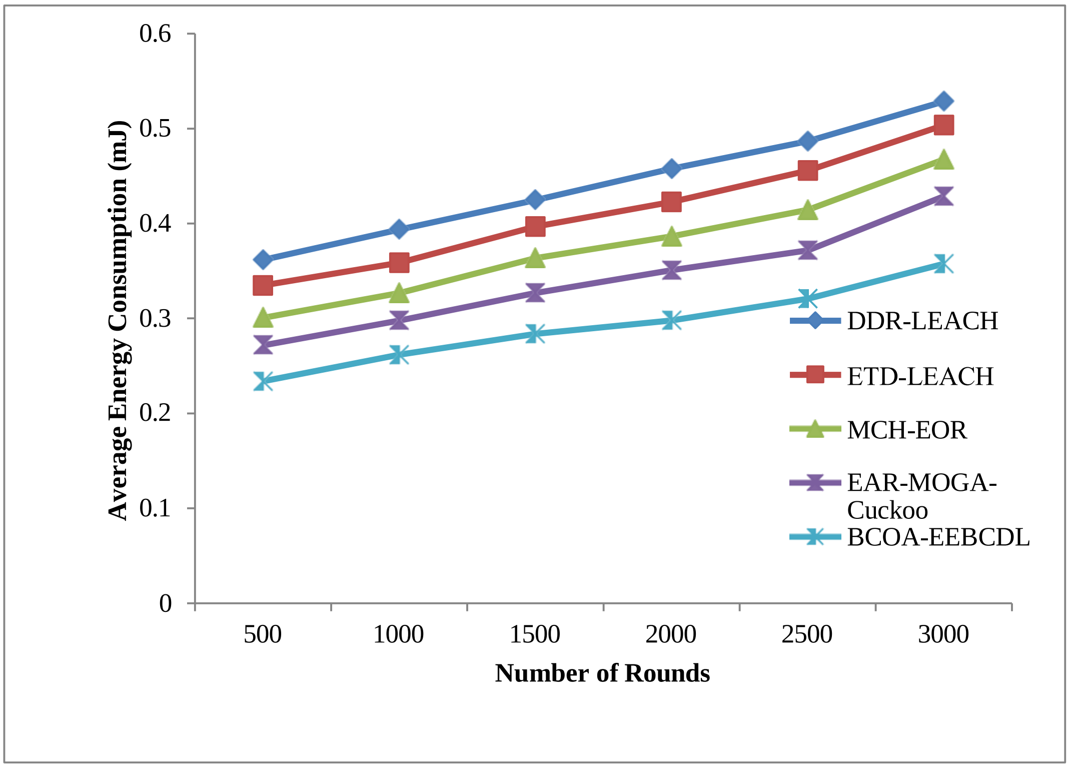

FIGURE 11. AVERAGE ENERGY COMPARISON VS. ROUTING PROTOCOLS

Figure 11 shows that the proposed method has very less average energy consumption compared with DDR-LEACH, ETD-LEACH, MCH-EOR, and EAR-MOGA-Cuckoo methods with a network size of 100 to 500 nodes. Proposed system has lowest average energy consumption of 0.358 mJ, other methods such as DDR-LEACH, ETD-LEACH, MCH-EOR, and EAR-MOGA-Cuckoo has increased average energy consumption of 0.529 mJ, 0.504 mJ, 0.468 mJ, and 0.429 mJ for 500 nodes (See Table 4).

5. CONCLUSION AND FUTURE WORK

Wireless Sensor Network (WSN) protocols are mainly used to achieve low power consumption of the network, but there are few studies on the quality of services (QoS) of these networks. In this paper, an Energy Efficiency Blockchain Deep Learning (EEBCDL) routing protocol with an inter-cluster routing in WSN. EEBCDL, optimal number of clusters k has been grouped by AGglomerative NESting (AGNES) with the network structure model and energy consumption model. Then, in order to satisfy the uniformity of the distribution of cluster heads (CHs), the monitoring area is divided into four small areas centered on the base station (BS), and the CHs are selected in the respective cells. Optimal CH is selected by the Beetle Chimp Optimization Algorithm (BCOA) with fitness function using constraints like residual energy, Distance among sensor nodes, CH and BS Distance, node degree, and node centrality. The residual energy is considered in the fitness function to avoid a dead node as a CH. Higher centrality, node degree is used to reduce the distance of transmission among the members. BCOA is enhanced with two steps: 1) Disruptive polynomial mutation is used to initialize the population, which provides the foundation for global search, 2) Next, reduce the probability of falling into the local optimum, the beetle antennae operator is used to improve the less fit chimps while gaining visual capability. Then, blockchain enabled RMCV and DL models, which detect the MNs. Quality routing is performed both with LEACH and HMGEAR only by the LNs found in the second layer. ANN, CNN, LSTM, and GRU models are trained on WSN-DS 2016 dataset generated at the first layer. Simultaneously, the trained DL models also detect the MNs in the network. Additionally, when MNs are removed from the network, the LNs are registered in the blockchain using PoA consensus mechanism. This work doesn’t optimize the information transmission path. Therefore, compared with some general low-latency protocols, its delay may be larger, which is not suitable for some projects with higher real-time requirements. Future will consider optimizing the information transmission path through a relatively practical optimization algorithm.

REFERENCES

1. Yahaya A. S., N. Javaid, M. U. Javed, A. Almogren, and A. Radwan, ‘‘Blockchain-based secure energy trading with mutual verifiable fairness in a smart community,’’ *IEEE Trans. Ind. Informat.*, vol. 18, no. 11, pp. 7412–7422, Nov. 2022.
2. Awan, S., Javaid, N., Ullah, S., Khan, A.U., Qamar, A.M. and Choi, J.G., 2022. Blockchain based secure routing and trust management in wireless sensor networks. Sensors, 22(2), pp.1-24.

3. Huang R., L. Ma, G. Zhai, J. He, X. Chu, and H. Yan, ‘‘Resilient routing mechanism for wireless sensor networks with deep learning link reliability prediction,’’ IEEE Access, vol. 8, pp. 64857–64872, 2020.
4. Elshrkawey, M., Elsherif, S.M., Wahed, M.E. (2018). An Enhancement Approach for Reducing the Energy Consumption in Wireless Sensor Networks, Journal of King Saud University - Computer and Information Sciences, 30(2), 259-267.
5. Bhushan, B., Sahoo, G. (2019). Routing Protocols in Wireless Sensor Networks, Computational Intelligence in Sensor Networks, 215-248.
6. Verma, P., Shaw, S., Mohanty, K., Richa, P., Sah, R., Mukherjee, A. (2018). A Survey on Hierarchical Based Routing Protocols for Wireless Sensor Network, 2018 International Conference on Communication, Computing and Internet of Things (IC3IoT), IEEE, 338- 341.
7. Zahedi, A.; Arghavani, M.; Parandin, F.; Arghavani, A. Energy Efficient Reservation-Based Cluster Head Selection in WSNs. Wirel. Pers. Commun. 2018, 100, 667–679.
8. Dutt, S.; Agrawal, S.; Vig, R. Cluster-Head Restricted Energy Efficient Protocol (CREEP) for Routing in Heterogeneous Wireless Sensor Networks. Wirel. Pers. Commun. 2018, 100, 1477–1497.
9. Neamatollahi, P.; Naghibzadeh, M. Distributed unequal clustering algorithm in large-scale wireless sensor networks using fuzzy logic. J. Supercomput. 2018, 74, 2329–2352.
10. Vijayalakshmi, K. and Anandan, P., 2019. A multi objective Tabu particle swarm optimization for effective cluster head selection in WSN. Cluster computing, 22(Suppl 5), pp.12275-12282.
11. Han, Y., Li, G., Xu, R., Su, J., Li, J. and Wen, G., 2020. Clustering the wireless sensor networks: a meta-heuristic approach. IEEE Access, 8, pp.214551-214564.
12. Khan, M.A.R., Shavkatovich, S.N., Nagpal, B., Kumar, A., Haq, M.A., Tharini, V.J., Karupusamy, S. and Alazzam, M.B., 2022. Optimizing hybrid metaheuristic algorithm with cluster head to improve performance metrics on the IoT. Theoretical Computer Science, 927, pp.87-97.
13. Sarkar, A. and Murugan, T.S., 2022. Analysis on dual algorithms for optimal cluster head selection in wireless sensor network. Evolutionary Intelligence, 15(2), pp.1471-1485.
14. Jia, H., Sun, K., Zhang, W. and Leng, X., 2021. An enhanced chimp optimization algorithm for continuous optimization domains. Complex & Intelligent Systems, pp.1-18.
15. Padmavathi U. and N. Rajagopalan, ‘‘Concept of blockchain technology and its emergence,’’ in Blockchain Applications in IoT Security. Hershey, PA, USA: IGI Global, 2021, pp. 1–20.
16. Abbas S., N. Javaid, A. Almogren, S. M. Gulfam, A. Ahmed, and A. Radwan, ‘‘Securing genetic algorithm enabled SDN routing for blockchain based Internet of Things,’’ IEEE Access, vol. 9, pp. 139739–139754, 2021.
17. Revanesh M. and V. Sridhar, ‘‘Hierarchical block chain-based authentication management scheme for IoT devices,’’ in Emerging Research in Computing, Information, Communication and Applications. Singapore:Springer, 2022, pp. 531–541.
18. Zhao, Z., Xu, K., Hui, G. and Hu, L., 2018. An energy-efficient clustering routing protocol for wireless sensor networks based on AGNES with balanced energy consumption optimization. Sensors, 18(11), pp.1-27.
19. Mehta, D. and Saxena, S., 2020. MCH-EOR: Multi-objective cluster head based energy-aware optimized routing algorithm in wireless sensor networks. Sustainable Computing: Informatics and Systems, 28, pp.1-34.
20. Baradaran, A.A. and Navi, K., 2020. HQCA-WSN: High-quality clustering algorithm and optimal cluster head selection using fuzzy logic in wireless sensor networks. Fuzzy Sets and Systems, 389, pp.114-144.
21. Khot, P.S. and Naik, U., 2021. Particle-water wave optimization for secure routing in wireless sensor network using cluster head selection. Wireless Personal Communications, 119(3), pp.2405-2429.
22. Adumbabu, I. and Selvakumar, K., 2022. Energy efficient routing and dynamic cluster head selection using enhanced optimization algorithms for wireless sensor networks. Energies, 15(21), pp.1-18.
23. Zhao, X., Zhong, W. and Navaei, Y.D., 2022. A Novel Energy-Aware Routing in Wireless Sensor Network Using Clustering Based on Combination of Multiobjective Genetic and Cuckoo Search Algorithm. Wireless Communications and Mobile Computing, 2022(1), pp.1-14.
24. Mistarihi, M.Z., Bany Salameh, H.A., Alsaadi, M.A., Beyca, O.F., Heilat, L. and Al-Shobaki, R., 2023. Energy-efficient bi-objective optimization based on the moth–flame algorithm for cluster head selection in a wireless sensor network. Processes, 11(2), pp.1-19.
25. Liu, Y., Huang, H. and Zhou, J., 2024. A Dual Cluster Head Hierarchical Routing Protocol for Wireless Sensor Networks Based On Hybrid Swarm Intelligence Optimization. IEEE Internet of Things Journal, IEEE Internet of Things Journal , vol. 11, no.9, pp. 16710 – 16721.

26. Amjad, S., Abbas, S., Abubaker, Z., Alsharif, M.H., Jahid, A. and Javaid, N., 2022. Blockchain based authentication and cluster head selection using DDR-LEACH in internet of sensor things. Sensors, 22(5), pp.1-20.
27. Khan, A.U., Sajid, M.B.E., Rauf, A., Saqib, M.N., Zaman, F. and Javaid, N., 2022. Exploiting Blockchain and RMCV-Based Malicious Node Detection in ETD-LEACH for Wireless Sensor Networks. Wireless Communications and Mobile Computing, 2022(1), pp.1-15.
28. Yang, J., He, S., Xu, Y., Chen, L. and Ren, J., 2019. A trusted routing scheme using blockchain and reinforcement learning for wireless sensor networks. Sensors, 19(4), pp.1-19.
29. Tian Y., Z. Wang, J. Xiong, and J. Ma, ‘‘A blockchain-based secure key management scheme with trustworthiness in DWSNs,’’ IEEE Trans. Ind. Informat., vol. 16, no. 9, pp. 6193–6202, Sep. 2020.
30. Daoud, M.S., Shehab, M., Abualigah, L., Alshinwan, M., Elaziz, M.A., Shambour, M.K.Y., Oliva, D., Alia, M.A. and Zitar, R.A., 2023. Recent Advances of Chimp Optimization Algorithm: Variants and Applications. Journal of Bionic Engineering, 20(6), pp.2840-2862.
31. Jia, H., Sun, K., Zhang, W. and Leng, X., 2021. An enhanced chimp optimization algorithm for continuous optimization domains. Complex & Intelligent Systems, pp.1-18.
32. Abed-alguni, B.H., 2019. Island-based cuckoo search with highly disruptive polynomial mutation. International Journal of Artificial Intelligence, 17(1), pp.57-82.
33. Chen, C., Cao, L., Chen, Y., Chen, B. and Yue, Y., 2024. A comprehensive survey of convergence analysis of beetle antennae search algorithm and its applications. Artificial Intelligence Review, 57(6), pp.1-72.
34. Lin, M., Li, Q., Wang, F. and Chen, D., 2020. An improved beetle antennae search algorithm and its application on economic load distribution of power system. IEEE Access, 8, pp.99624-99632.
35. Kurani, A., Doshi, P., Vakharia, A. and Shah, M., 2023. A comprehensive comparative study of artificial neural network (ANN) and support vector machines (SVM) on stock forecasting. Annals of Data Science, 10(1), pp.183-208.
36. Manzoor, I. and Kumar, N., 2017. A feature reduced intrusion detection system using ANN classifier. Expert Systems with Applications, 88, pp.249-257.
37. Li, Z., Liu, F., Yang, W., Peng, S. and Zhou, J., 2021. A survey of convolutional neural networks: analysis, applications, and prospects. IEEE transactions on neural networks and learning systems, 33(12), pp.6999-7019.
38. Purwono, P., Ma'arif, A., Rahmaniar, W., Fathurrahman, H.I.K., Frisky, A.Z.K. and ul Haq, Q.M., 2022. Understanding of Convolutional Neural Network (CNN): A review. International Journal of Robotics and Control Systems, 2(4), pp.739-748.
39. Livieris, I.E., Pintelas, E. and Pintelas, P., 2020. A CNN–LSTM model for gold price time-series forecasting. Neural computing and applications, 32, pp.17351-17360.
40. Wen, X. and Li, W., 2023. Time series prediction based on LSTM-attention-LSTM model. IEEE Access, 11, pp.48322-48331.
41. Hussain, B., Afzal, M.K., Ahmad, S. and Mostafa, A.M., 2021. Intelligent traffic flow prediction using optimized GRU model. IEEE Access, 9, pp.100736-100746.
42. Jia, P., Liu, H., Wang, S. and Wang, P., 2020. Research on a mine gas concentration forecasting model based on a GRU network. IEEE Access, 8, pp.38023-38031.
43. Deng, X., Li, J., Ma, C., Wei, K., Shi, L., Ding, M., Chen, W. and Poor, H.V., 2022. Blockchain assisted federated learning over wireless channels: Dynamic resource allocation and client scheduling. IEEE Transactions on Wireless Communications, 22(5), pp.3537-3553.