

Energy Efficient Data Distribution on Cloud with Optimal Routing Path Based Congestion Control in WSN Environment

Pampapathi B M^{1*}, Nageswara Guptha M², M S Hema³

^{*1}Research Scholar, Department of Computer Science & Engineering, Sri Venkateshwara College of Engineering, Bengaluru, Karnataka, INDIA, Email id: bm.pampapathi77@gmail.com

²Professor, Department of Computer Science & Engineering, Sri Venkateshwara College of Engineering, Bengaluru, Karnataka, INDIA, Email id: mnguptha@yahoo.com

³Associate Professor, Department of Information Technology, Anurag University, Hyderabad, Telangana, INDIA, Email id: ghema_shri@yahoo.co.in

Abstract: Presently, Wireless Sensor Network (WSN) is the extremely standard services utilized in commercial in addition to industrial applications, owing to its technical development in the embedded computing devices' processor, communication, along with less-power usage. Moreover, in the WSN's advantages, the sensor data's energy-efficient distribution onto the cloud data centre is amid the core challenges caused by higher data loads and also congestion. Aimed at overcoming these problems, this work proffers an energy-efficient Data Packet (DP) distribution protocol comprising an optimal routing path centred congestion control technique in the WSN environment. This system not just ponders the energy-efficient DP distribution nevertheless increments the sensor's lifetime by choosing the Cluster Head (CH) centred on the sensor's residual energy, packet loss factor, data forwarding rate, and minimal distance. Next, the numerous routing paths prevalent in inter-cluster and intra-clusters are detected utilizing the QLA2ODV method; the congested nodes are detected as of these paths utilizing LM-NN classifier; if congestion happens, the QLA2ODV distributes the DP onto the Base Station (BS) via the optimum congestion-free routing path. The optimal routing paths were elected as of the congestion-free paths utilizing the WFCSO method. Past obtaining the sensor data, the BS transmits the DPs onto the Fog Nodes (FNs). At last, the sensor data are effectively distributed onto the cloud data centres grounded on the data's features and cloud features utilizing the BD-SBO technique. The experiential outcomes determine the proposed system's effectiveness.

Keywords: Wireless Sensor Network (WSN), Energy Efficient Data Distribution, Cloud Data Centers, Queue Load Adopted Ad-Hoc on-demand Distance Vector (QLA²ODV) protocol, Optimal Path, and Congestion detection.

1. INTRODUCTION

Enormous quantity of distributed Sensor Nodes (SN) that transmits information to the Cloud Server is possessed by IoT [1]. The SN introduced '3' major operations: (i) data collector, (ii) data processing, along with (iii) data transmission [2]. Industrial monitoring along with control systems utilized WSN for improving their functionality and also efficiency [3]. These networks seek many attractive applications in fields, namely business, healthcare, military surveillance, air pollution control, river level variation monitoring, intelligent highway designing along with remote health assistance, etc. For communication, data processing, along with perception, enormous energy is utilized by the SN while installing such applications [4]. The major performance metrics are network lifetime along with quick data delivery [5] because generally, SN has low power along with the battery's replacement is complicated [6].

Communication overhead is brought about by the strategy in which each SN communicated straight with the BS. For avoiding this, SN is categorized into disparate clusters [7]. One header is possessed by each cluster of a network, called a CH [8] that communicates with other CH on the network. CH's function is to gather the information detected from the SN that are the members in the clusters. The data is aggregated and sent to the BS or through the CH [9]. The CH's selection is executed in such a way that to attain higher throughput centered on the node's energy present within the cluster [10]. A lot of energy is required for converting perceived data directly to BS, and a routing protocol is utilized in cluster WSNs for identifying the better way between CH and PS to decrease power consumption [11].

A huge quantity of data transferred has brought dynamic modifications in the multiple routing paths, along with congestion is created within the other SN [12]. Therefore, no data is acquired by the BS from the SN [13]. Congestion stands as an undesirable occurrence owing to the significant reduction in network's

performance and network life and less energy wastage. Thus, greater consideration is gained by CC in WSN recently [14]. The best mitigation technique is to choose the best possible route for avoiding congestion. A process of suitable path selection as of source to destination for sending the DPs is routing [15]. Mathematical optimization, along with heuristics techniques, is utilized to conduct the route discovery method [16]. Numerous nodes implicated in route discovery are reduced by these techniques, which will reduce the network's control overhead and also the system's Energy Consumption (EC) [17].

Apart from the above-mentioned problems, energy-effective distribution of sensor data on virtual machines (VM) has seen as an essential problem on the cloud computing environment by considering several elements like completion time, the all-out expense for conducting every one of clients' assignments, asset's usage, power utilization, along with adaptation to internal failure [18]. The overall network for optimal path discovery and node's EC must be controlled for designing an energy-effective data transmission of sensor data on VM. The routing protocol needs to be adequately effective for sustaining the throughput, routing overhead reduction, packet delivery ratios (PDR), together with the decrementation of Ends-to-Ends Delay (EED) in the complete network [19]. Thus, an energy-effective data distribution scheme with an optimal routing path centered CC approach in a WSN environment is proposed by this paper.

This paper is categorized as: Section 2 explained the related works to the proposed CC and energy-effective data distribution. Section 3 presented the proposed system's accurate overview. Section 4 examined and assessed the proposed system's outcomes. Section 5 completed the paper together with future guidelines is provided.

2. RELATED WORKS

Pampapathi B M *et al.* [1] introduced 'two' methods for IoT data distribution to the cloud utilizing IANFIS in addition to secure transmission of those distributed data utilizing MECC. Initially, the IoT device is authenticated by means of performing registration, login, as well as verification for preventing the device as of unauthorized utilization. Subsequently, the device is accessed to sense the information as of associated objects by utilizing smart sensors; in addition, the communication of disparate IoT sensors is performed by employing the MQTT protocol. By means of the IANFIS, this sensed information as of disparate IoT devices is distributed to the specific cloud server (CS). Then by employing the MECC algorithm, the data is sent securely to the user. Lastly, the local calculation (LC) is carried out by means of HFPGA. The experimentation is executed for analyzing the proposed work's performance. The outcomes exhibited that the proposed techniques work better when weighed against other existing algorithms.

D. Loganathan *et al.* [20] created an Energy-Aware Efficient Data Aggregation (EAEDAR) as well as Data Re-Scheduling with the amalgamation of clustering methods. This model was formed with an energy factor-centered cluster for reducing the transfer time along with improving the data transmission efficiency. The SN's EC rate was lessened by implementing the Appropriate Aggregator SN selection along with effective Data Aggregation, thus enhancing the network's life. Data PL was reduced by utilizing Transmission Re-scheduling. The EAEDAR model had provided a higher PDR, throughput, along with energy efficiency as shown by the results. But, the transmission delay was not decreased to the normal level.

MortezaKheirkhah *et al.* [21] introduced a transport protocol, i.e. Maximum MultiPath TCP (MMPTCP), for minimizing short flows' completion times whilst producing outstanding goodput to long flows. MMPTCP operated in '2' stages. First, the packets were randomly scattered across the network in a congested window. This was beneficial for latency-sensitive flows. MMPTCP had switched to normal Multiple Path TCP modes after a particular data amount was sent. The number of subflows was adjusted by a topology-specific extension of MMPTCP during the protocol's second stage centered upon knowledge of the recipient's location in the data center. MMPTCP had substantially decreased the loss rate of every link within the network along with network resources' utilization was increased as shown by the results.

Baradaran and Navi [22] presented a Higher-Quality Clustering Algorithms (HQCA) to generate higher-quality clusters for providing energy-effective data transmission. A criterion was utilized by the HQCA technique for assessing the cluster quality that improved the inter-cluster along with intra-cluster distances in addition to error rate during clustering was reduced. The optimal CH was chosen centered on fuzzy logic; as per several criteria, namely the residual energy, the minimum and maximum energy in each cluster, the minimum and also maximum distances among the nodes within all clusters and also BS. The HQCA-WSN had improved

the network's lifetime, even though the average EC in HQCA-WSA must be reduced as shown by the simulation results.

Shishupal Kumar *et al.* [23] introduced a G.9959 protocol-centered packet delivery method in a WSN environment. The encoding along with decoding was performed on the network sensed packet parameters intended for selecting an appropriate channel/path to support multicasting. The presented strategy was utilized to assess the IoT-oriented WSN's performance concerning several performance metrics, say energy, lifetime, overhead along with PDR, etc. This scheme had outperformed when contrasted to other approaches predicted in this field as shown by the simulation outcomes. A higher probability was given for effectively transmitting the packet to the destination with EC. A transmission delay was caused by this approach because of congestion occurrence.

Mehta and Saxena [25] introduced a Multiple Objective centered Clustering along with Sailfish Optimizers (SFO) technique for sustaining WSN's energy efficiency. The CH was selected centered upon efficient fitness function that was created from many objectives. It had assisted to diminish the consumption of energy together with various dead SN was reduced. SFO was utilized for selecting an optimum path to S_n intended for transmitting data after the CH selection. The SFO method had performed 21.9% and 24.4% better concerning EC along with many live SN respectively when contrasted to GWO as shown by the simulation outcomes. The network's lifetime will be considerably decreased if there is more traffic movement from several CH to the sink.

Amit Singh and Nagaraju [26] introduced an application of '3' various algorithms for improving the sensor network's performance. Moreover, opportunistic coding was utilized for reducing several transmissions at potential relays. Therefore, '3' algorithms were integrated by this implementation that combined the merits of each for considerable improvement in data transmission. Firstly, the S_n was placed through a particle swarms optimizations. Secondly, the route construction from the sensors along with a sink of the particular cluster utilizing a minimum wiener spanning tree was further optimized by the artificial bee colony method. Opportunistic packet combination was done before sending to neighbors in the third step. An enhancement of 25% network throughput was attained as shown by the results. Higher time complexity was the major drawback in this method.

Pampapathi B M *et al.* [28] proposed a secure and energy-aware routing method of WSN using Cosine Distance-based Learning Vectors Quantization (CD-LVQ) and Brownian movement-based Shark Smell Optimizations (BM-SSO) algorithms. Initially, the classifier, namely CD-LVQ, finds the MN on the WSN. It extracts some node parameters for detection. After that, only Trusted Nodes (TN) in the networks is considered for further operation by avoiding those MN. The GRBF-MS clustering algorithm groups the TN; for each grouping of SN, the Cluster Head (CH) is selected to increase the Network Life-Time (NLT) and save the nodes' energy. Next, the BMSSO algorithm selects an optimal path for the CH that generates the best path as of CH to sink aimed at Data Transmission (DT). Lastly, the Ranking mutation-based Blue Monkey Optimization (RMBMO) efficiently allocates the sink data to the optimal Virtual Machine (VM) of the Cloud Server (CS). The proposed methods' outcomes are examined and contrasted with the other prevailing techniques, which confirmed that the proposed CD-LVQ, BMSSO, and RMBMO algorithms classify the MN, selecting the optimal path, and perform rescheduling in an extremely good manner when compared with others.

3. CONGESTION AVOIDANCE AND ENERGY EFFICIENT DATA DISTRIBUTION

Recently, wireless SN operates as a base aimed at the higher-end industrial Internet of Things (IoT) application. All real-time IoT application senses data employing previously built-in sensors. Sensor DPs' energy-efficient distribution on VMs is amid the core challenges prevalent in WSN. Networks can be disturbed by congestion caused by the network's limited resources and its higher data loads. Congestion and unbalanced EC prevalent in the DPs' routing procedure is amidst the core problems existent in WSN. A proper method must be adapted aimed at modelling the energy-efficient data distribution system comprising congestion control in the WSN that evades DP loss, decrements the power consumption, and also increments the network life. Consequently, an energy-efficient DP distribution technique comprising optimal routing path centred congestion control methodology put forwarded in this work has been modelled aimed at WSN. The proposed system comprises '6' stages: inter-cluster and intra-cluster formation, numerous routing path creations, congestion detection, optimal path centred congestion evasion, energy-efficient data distribution over the VM. The stages are detailed below. Figure 1 exhibits the proposed congestion control system's flow.

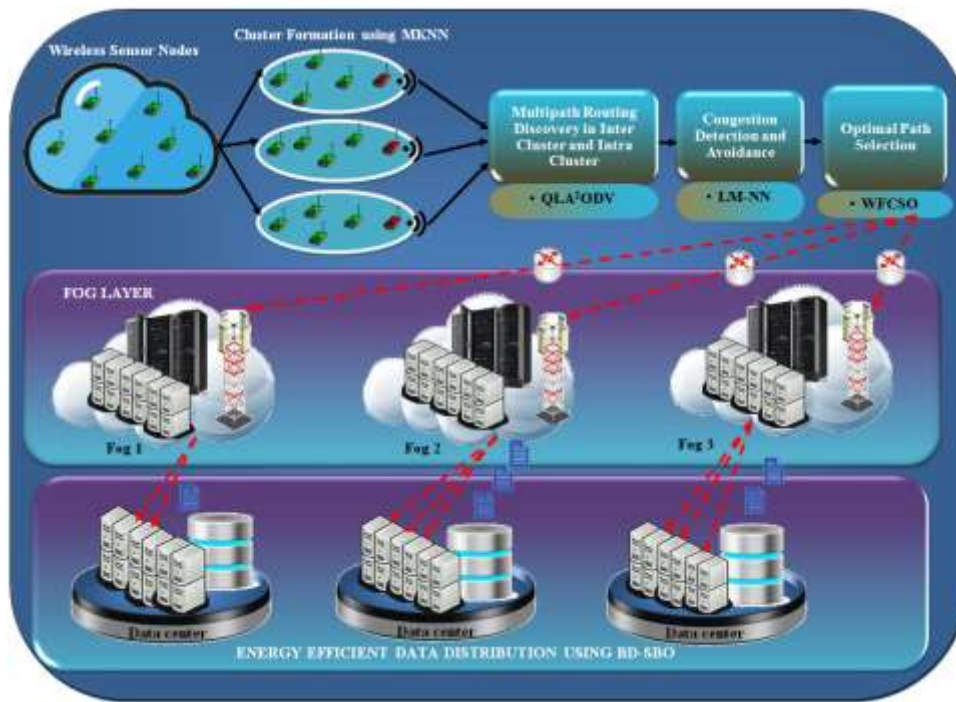


Figure 1: Block Diagram of Proposed Methodology

3.1 Cluster Formation

The SNs are the minute autonomous devices that are set up in the WSN's sensing area. They sense the natural events as of the sensing domain and then send the sensed data towards the BS via multiple-hop communication. At first, the SNs will be clustered utilizing Manhattan distance-centred K-Nearest Neighbour (MKNN) clustering protocol. To resolve the energy issues, a CH is chosen here, where the source node (SN) sends its data onto the CH via the intermediate nodes within every cluster. K-Nearest Neighbour is amidst the simplest Machine Learning protocol grounded on the Supervised Learning methodology. It ponders the similarity betwixt the novel data and existent DP and adds the novel data underneath the group which is much identical to the prevalent types. Aimed at computing the similarity betwixt '2' data, the proposed MKNN technique utilizes Manhattan distance that is computed as the addition of the absolute differences betwixt the '2' vectors such that the cluster formation is much precise. The MKNN cluster the SNs grounded on their distance as of their neighbouring nodes. The MKNN's steps are given as,

Step 1: Initialize the L -number of clusters. Ponder $n_i = \{n_1, n_2, \dots, n_N\}$ as the set of SNs and $p_i = \{p_1, p_2, \dots, p_N\}$ as the SNs' position.

Step 2: Elect m_i random SNs as neighbours. Aimed at every SN n_i perform the subsequent functions:

- (a) Compute a N number of neighbours' Manhattan distance as,

$$MT_{dist}(n_i, m_i) = \sum_{i=1}^N |(n_i - m_i)| \quad (1)$$

- (b) Organise the distances' and indices' ordered collection as of the smallest to the largest (in the ascending order) via the distances.
- (c) Elect the 1st N entries as of the organised collection
- (d) Amidst the N -neighbours, calculate the number of data-points within every group.
- (e) Then, it will allot a class for the SN centred on the largely frequent class of these rows.

Step 3: Continue step 2 till the identical points are allotted to every cluster within the successive rounds. Pursuing this method, every node prevalent in WSN is clustered.

Past the cluster formation, the CH aimed at every respective cluster is chosen grounded on the sensor's residual energy, packet loss factor, minimal distance, and also data forwarding rate that is fixed as the fitness function. Ponder, $J_i(E_{res}, D_f, D_t, P_{lr})$ as the i -th SN's residual energy, data forwarding rate, distance, and packet loss factor. The proposed protocol encompasses '1' preset threshold value that the SN's $J_i(E_{res}, D_f, D_t, P_{lr})$ value nearest to the threshold is pondered as a CH; likewise, the CHs are elected aimed at every cluster. These CHs are accountable to gather the DP as of the SN.

3.2 Multiple Routing Discovery in Inter-Cluster and Intra Cluster

Many alternate paths through a network are created by a routing method named Multipath routing discovery. Several benefits, namely augmented bandwidth, fault tolerance, along with enhanced security are produced. The single routing path connecting the source as well as the destination node (DN) is found by many WSN routing protocols. When the primary path breaks or congestion appears, the DP are dropped by the intermediate nodes as alternate paths are unavailable to arrive at the DN till the subsequent route discovery method is started. Thus, multiple routing paths betwixt nodes are created by the proposed work inside the cluster, which are named inter-cluster paths. Moreover, the proposed work produces multiple routing paths between CH, which are labeled intra-cluster paths. The proposed system utilized the Queue Load Adopted Ad-Hoc on-demand Distance Vector (QLA²ODV) protocol for the discovery of every possible path to the destination.

Route request (r_RQ) messages are transmitted within the network by QLA²ODV for discovering a possible route to the DN. The information regarding the route to the source nodes, previous nodes, along with forwarding nodes in the route is maintained by intermediate nodes on route reply (r_RP) message from the DN to the source. Conventionally, routing tables (RT) are utilized by the Ad-Hoc on-demand Distance Vector protocol for storing routing information. The RT includes the subsequent fields (columns): source address, source sequence number, broadcast id, destination sequence number, a destination address, hop_count, and life_time. However, another column i.e. queue load is included by this proposed methodology. Now, the mobile agent is initialized for monitoring and updating the queue load of all the nodes on the QLA²ODV routing table for congestion path detection. Hence, the mobile agent is merged with the QLA²ODV protocol in the proposed system.

Table 1: Routing Table Format of QLA²ODV protocol

Source Address	Source Sequence Number	Broadcast_id	Destination Address	Destination Sequence Number	Hop_Count	Queue Load
----------------	------------------------	--------------	---------------------	-----------------------------	-----------	------------

The RT is maintained at each node. The discovered multiple routing paths are mathematically indicated as,

$$r_i = \{r_1, r_2, r_3, \dots, r_n\} \quad 1 \quad (2)$$

Where, r_i represents the discovered routing paths as well as r_n represents n -th routing path. The DP is transmitted to the DN by means of the source through these routing paths.

3.3 Congestion Detection and Avoidance

Herein, every node's status is calculated to decide the quantity of traffic prevalent across diverse routing paths. Congestion on a network occurs at any time; whilst the number of DPs arriving at a node surpasses its buffer capacity, the node gets congested and starts to lose DPs. Hence, it is essential to detect every node's

congestion status within the routing paths in advance. Centred on every node's queue load, the congestion may be found. A node's queue load signifies the number of DPs prevalent within the queue at any specified time t . Whilst the queue's load rises, congestion occurs. Whilst the queue's load value is big, the arriving traffic gets bursty. The queue load's continuous growth signifies that the arriving heavy traffic surpasses the mobile node's buffer capacity and also buffer overflow happens.

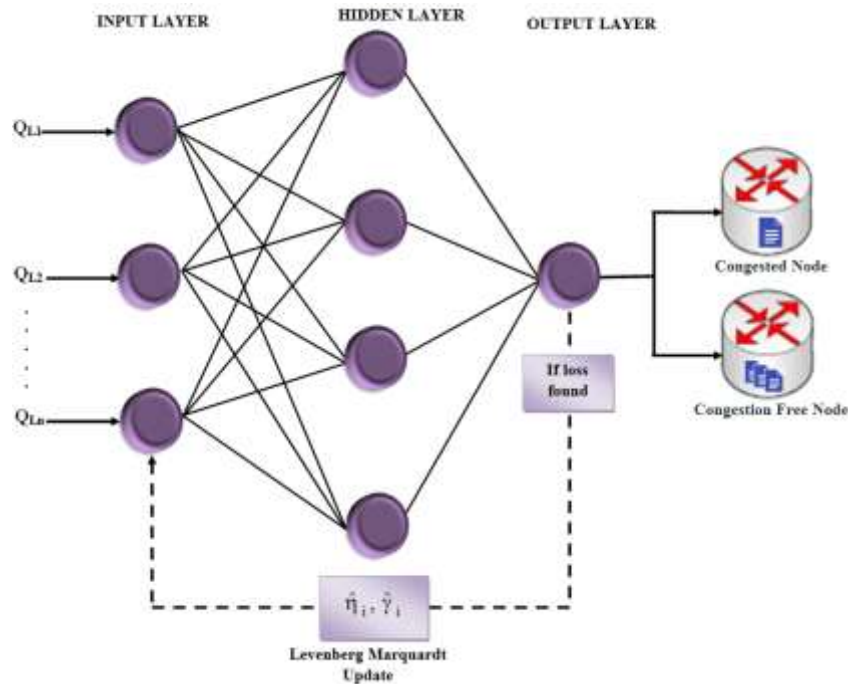


Figure 2: Structure of LM-NN

Aimed at congestion detection, the proposed system utilizes the LM-NN that ponders every node's queue load as input and also detects the congestion in advance. Amid the very crucial factors influencing the system's performance is the training function's selection that is changed grounded on the system's loss computation. Past loss computation, the bias value and Weight Value (WV) are updated grounded on the Levenberg-Marquardt since it is to detect efficient training directions utilizing the loss function's 2nd derivatives. Figure 2 shows the LM-NN's general structure.

Aimed at congestion identification, the queue's load values have been fed initially into the inputted layer neurons. Each node's queue load is articulated as,

$$Q_{Li} = \{Q_{L1}, Q_{L2}, Q_{L3}, \dots, Q_{Ln}\} \quad (3)$$

Here, Q_{Li} implies the i -th node's queue load. Every node's queue load is enumerated as,

$$Q_L = Q_{inst} - Q_{avg} \quad (4)$$

Here, Q_{inst} signifies the instant queue size; Q_{avg} implies the average queue size that is enumerated as,

$$Q_{avg}(t) = (1 - \bar{w}) * Q_{avg}(t-1) + Q_{inst} * \bar{w} \quad (5)$$

Here, \bar{w} signifies the queue's weight that is a constant parameter ($\bar{w} = 0.02$). Offering the nodes' queue load Q_{Li} into the inputted layer neurons, an equivalent WV η_i is initialized. Then both the inputted valued and also WVs are offered to the hidden layer that generates the subsequent output.

$$\alpha_h = \gamma + \sum_{i=1}^n Q_{Li} * \eta_i \quad (6)$$

Here, α_h signifies the hidden layer's output, γ implies the bias value. After that, the end output is enumerated by summing all input values' weights.

$$\alpha_o = \gamma + \sum_{i=1}^n \alpha_H * \eta_i \quad (7)$$

Here, α_o implies the outputting unit. Next, the network training algorithm is centred on an energy function's minimization signifying the instant error.

$$E_i = \hat{\alpha}_o - \alpha_o \quad (8)$$

Here, E_i implies the error function, α_o signifies the attained output; $\hat{\alpha}_o$ signifies the NN's aimed output. The system's error is should be decremented. Aimed at decrementing this value, the bias and weights value should be updated. In this training procedure, the Levenberg-Marquard is utilized aimed at updating bias value and also WV. The WV's and bias value's updation procedure utilizing Levenberg-Marquard is equated as,

$$\hat{\eta}_i = \eta_i - (J_i^T * J_i + \delta_i I) \cdot (2J_i^T \cdot E_i) \quad (9)$$

$$\hat{\gamma}_i = \lambda_i - (J_i^T * J_i + \delta_i I) \cdot (2J_i^T \cdot E_i) \quad (10)$$

Here, $\hat{\eta}_i$ implies the updated WV; $\hat{\gamma}_i$ implies the bias value; δ_i implies the damping factor, J_i symbolizes the Jacobian matrix. The updated bias values and WVs are dispersed to the LM-NN. The bias values and WVs are updated until attaining the minimal error value. Thus, the LM-NN has been trained to detect the congestion prevalent in the network.

The LM-NN's output comprises '2' classes: congestion-free nodes, congested nodes. If a congested node is offered, the paths that comprise that congested nodes are evaded aimed at the upcoming data packet transmission till the normal traffic condition is reached. The routing protocol QLA²ODV begins to distribute the traffic over other congestion-free paths for attaining the normal traffic condition to evade congestion.

3.4 Optimal Routing Path Discovery

Aimed at attaining minimal transmission delay, least EC, together with maximal throughput, the routing paths must be consistent, greatly energy-efficient, and shortest as of the source node to the DN. Consequently, the optimal routing path is identified utilizing the Wiener Factor Chicken Swarm Optimization (WFCSO). WFCSO imitates a chicken swarm's hierarchical order together with its behaviours at the food search, in which every chicken signifies a potential solution for an optimization issue. Basically, CSO employs the following '4' rules for idealizing the chickens' behaviour:

1. The chicken flock comprises numerous groups. Every group encompasses '1' dominant rooster, '2' chickens, and then chicks.
2. The Chickens comprising the finest multiple fitness values act as roosters, each one will be the head rooster in their group. Chickens comprising the least multiple exercise values will be hired as chicks and the others will be chickens.
3. The dominant relationship and the mother-child relationship stay the same in the group hierarchical order. These classes just update every number (G) step.
4. Chickens pursue their rooster mate whilst seeking food. Ponder that the chickens generally steal the good quality food which others have found already. Chicks seek food just in the region of their mothers. The rooster is in support of the competition aimed at food.

Chickens comprising good fitness values prioritize food access analogized to least fitness values. Likewise, if a path comprises higher residual energy and also the shortest path distance, that path is more likely to be optimal. This function is preset as WFCSO's fitness function. Each solution's fitness is computed; every candidate solution is pondered as the elected path. Consequently, each solution's fitness is pondered as the least EC and least path distance.

$$F_i(opm) = \begin{cases} \max(E_{res}); & \text{Maximum residual energy} \\ \min(dist(n_s, n_d)); & \text{Minimum path distance} \end{cases} \quad (11)$$

Here, $dist(n_s, n_d)$ signifies the path distance betwixt a source node and a DN. Centred on the fitness values, the chicken swarm is split as diverse subgroups. Ponder that E_R, E_H, E_C and E_M as the number of roosters, the hens, the chicks and the mother hens, correspondingly. All n chicken is signified by its position as,

$$E_{i,j}(t) = \{E(t)_{1,1}, E(t)_{1,2}, E(t)_{1,3}, \dots, E(t)_{n,d}\} \quad (12)$$

Here, $E_{i,j}(t)$ implies the n chickens' position at iteration t that look for food within the d -dimensional space. Since the chickens' behaviours change inside the subgroup, the rules aimed at updating their status also change regarding their roles.

Rooster Movement: Rooster comprising good fitness values finds a wider range of food analogized to those comprising least fitness values. Their movement is equated as,

$$E_{i,j}(t+1) = E_{i,j}(t) * (1 + \omega_f(r)); \quad (13)$$

$$\kappa^2 = \begin{cases} 1, & \text{if } F_i \leq F_v \\ \exp\left(\frac{F_v - F_i}{|F_i + \xi|}\right), & \text{Otherwise} \end{cases} \quad v \in [1, n], v \neq i \quad (14)$$

Here, ξ signifies the smallest constant utilized to evade the zero-division-error; v implies a rooster's index that is chosen randomly as of the roosters' group; F_i symbolizes the rooster E_i 's fitness value; $\omega_f(\bullet)$ signifies the wiener factor that is computed as,

$$\omega_f(r) = \frac{1}{\sqrt{n}} \sum r \quad r \in [0, \kappa^2] \quad (15)$$

Here, r implies the random variables.

Hen Movement: Regarding the chickens, they can pursue their group-corresponding roosters in food search. They will generally snatch the good quality food discovered by the other chickens, even if they are covered up by the other chickens. The hen movement's arithmetical design can be written as follows,

$$E_{i,j}(t+1) = E_{i,j}(t) + B_1 * \omega_f(r) * (E_{R_1,j}(t) - E_{i,j}(t)) + B_2 * \omega_f(r) * (E_{R_2,j}(t) - E_{i,j}(t)) \quad (16)$$

$$B_1 = \exp\left(\frac{f(i) - f(r_1)}{|f(i)| + \xi}\right) \quad (17)$$

$$B_2 = \exp(F_{R_2} - F_i) \quad (18)$$

Here, $r \in [0,1]$, R_1 implies the rooster's index that is the i^{th} hen's group-mate, whilst R_2 implies a chicken's index (hen or rooster) that is chosen randomly as of the swarm $R_1 \neq R_2$.

Chick Movement: Chicks search just around their mother for food. The chick movement's arithmetical design is proffered in eqn. (),

$$E_{i,j}(t+1) = E_{i,j}(t) + \omega_f(Y) * (E_{M,j}(t) - E_{i,j}(t)) \quad Y \in [0,2] \quad (19)$$

Here, $E_{M,j}(t)$ signifies the position of the i^{th} chick's mother ($M \in [1, n]$); Y implies the parameter, which signifies how much fast a chick will pursue its mother. Examine each new solution's fitness. If the novel

solution's fitness is good whilst analogized to its preceding one, it is pondered as an optimal solution. Stop if the termination standard is fulfilled, or else execute the '3' steps above. Figure 3 exhibits the WFCSO algorithm's pseudo-code.

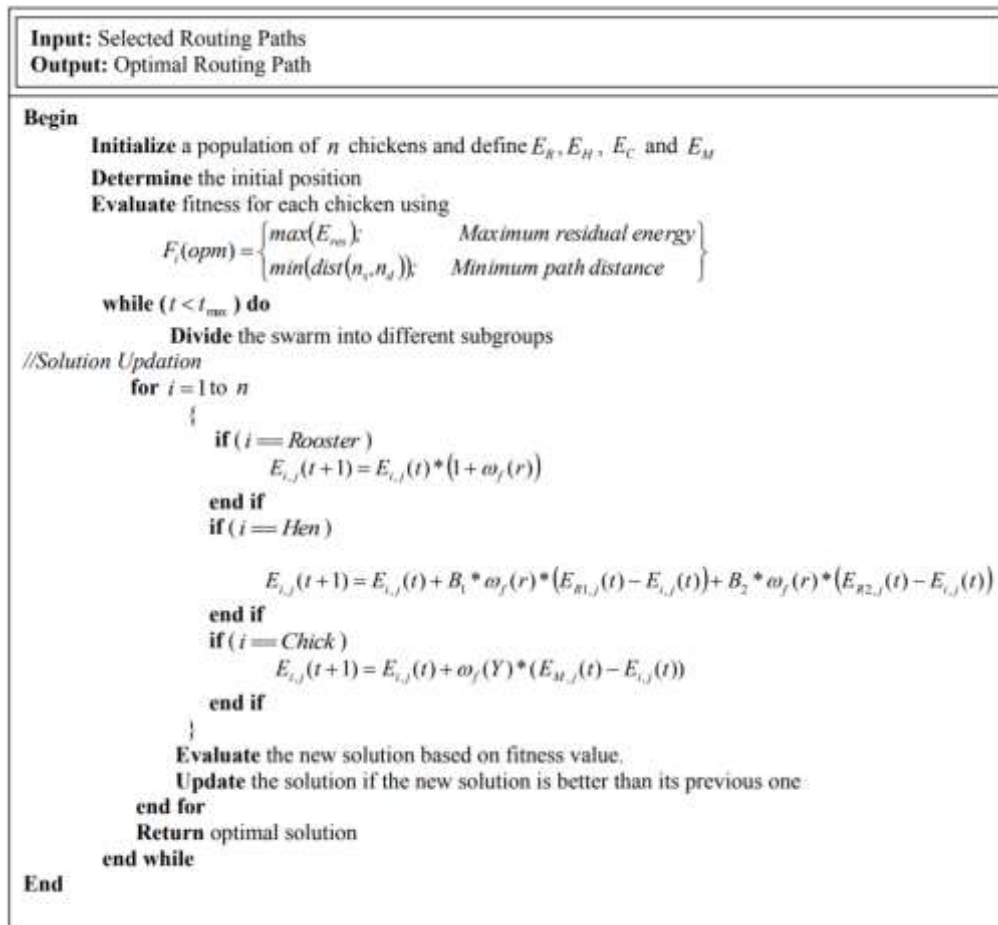


Figure 3: Pseudocode of WFCSO Algorithm

Likewise, the optimal path is elected by implementing this WFCSO algorithm betwixt the source and destination. The optimal path signifies the path, which comprises the highest energy-level with minimal distance. On this optimal routing path, the DPs are sent as of the SN to the DN; then, it is sent towards the BS. Past receiving of DP as of the wireless SNs, the BS combines the DP and transmits it towards the FNs (fog layer), which are the distributed fog calculating entities that facilitate the fog services' deployment; FNs are built by numerous physical devices comprising processing and sensing abilities. These nodes distribute the acquired sensor data into the cloud data centre with the objective to increment the network's bandwidth, privacy, mobility, and security.

3.5 Energy-Efficient Data Distribution

The process of organizing the resources among disparate VM as per certain rules and regulations of resource usage under a specific cloud environment is indicated by resource allocation. The SN data is distributed by the proposed work to the VM as per both node features along with VM features, say data type, data memory utilization, storage, processing speed, and bandwidth. Bernoulli Distribution (BD) adoptive Satin Bower-bird Optimizer (BD-SBO) was utilized for employing the resource allocation. The normal distribution is employed by the SBO in the bowerbird attraction phase, although it is inappropriate for a huge population. Therefore, BD is utilized here that performed superior to the normal distribution for a huge amount of values. The optimization process of BD-SBO is separated into four stages, which are described below.

An initial populace is produced randomly. This population is a set of series bowers, that is written as,

$$f_i = \{(f_1, f_2, \dots, f_N)_{d_1}, \dots, (f_1, f_2, \dots, f_N)_{d_N}; list(bowers)\} \quad (20)$$

Where, f_i indicates the set of bowers. The male bird is the important factor in building the attraction probability of the bowers whether it can attract a mate. A bower is selected by the female Satin bower bird centered upon its probability. The probability (ρ) can well be exhibited as,

$$\rho = \frac{F(i)}{\sum_{i=1}^{d_N} F(i)} \quad (21)$$

Where, $F(i)$ indicates the fitness of i -th solution, and d_N signifies the total bowers. The fitness value of each bower is computed by,

$$F(i) = \begin{cases} \frac{1}{1 + O(f_{d_N})}; & O(f_{d_N}) \geq 0 \\ 1 + |O(f_{d_N})|; & O(f_{d_N}) < 0 \end{cases} \quad (22)$$

Where, $O(f_{d_N})$ indicates the objective function value of d_N -th bower. The highest fitness value bower is deemed elite centered on the fitness value of each bower. The changes of every bird's bower in each iteration formulated as,

$$H_{d_N}(t+1) = H_{d_N}(t) + v_{d_N} \left(\left(\frac{H_{sd_N} + H_{elite,d_N}}{2} \right) - H_{d_N}(t) \right) \quad (23)$$

Where, $H_{d_N}(t+1)$ symbolizes the position of d_N -th bower at iteration $(t+1)$, H_{sd_N} indicates the larger probability having a solution, H_{elite,d_N} denotes the location of the elite, and v_{d_N} is the attractiveness parameter which determines as follows,

$$v_{d_N} = \frac{\lambda}{1 + \rho_k} \quad (24)$$

Where, λ represents the greatest step size. When the target position's probability is 0, the highest step size occurs while the step size will be λ . Conversely, when the target position's probability is 1, the lowest step size occurs while the step size will be $\lambda/2$. Next, the mutation is executed. A bower built on the ground by the males could be attacked by other animals or entirely ignored. The stronger males might steal items as of weak males or demolish their bowers in multiple cases. Hence, random changes with a BD and certain probability are applied in the mutation process.

$$H_{d_N}(t+1) \sim B(H_{d_N}(t+1), \sigma^2) \quad (25)$$

Where, $B(\bullet)$ indicates Bernoulli distributions, σ is a space width proportion, which is formulated as follows,

$$\sigma = D * (ub - lb) \quad (26)$$

Where, ub and lb denotes the upper as well as lower bound and D denotes the percentage of the difference betwixt the upper and lower bound. The old and new populations are combined and sorted centered on their fitness at the end of the process along with the new population is produced. When the termination conditions are not fulfilled, the bower position updation and mutation process are repeated. The optimal solution is finally obtained. Following this way, the VM is chosen for each SN data by applying this BD-SBO algorithm. Lastly, the SN data are efficiently distributed over the VM.

4. RESULTS AND DISCUSSIONS

The proposed system's performance is examined. This work is performed in the working platform of JAVA. The proposed system's performance is examined in '3' methods. Firstly, the performance analysis of optimal path selection is performed in section 4.1. Then, the performance analysis of congestion detection is executed in section 4.2. The performance analysis of energy-efficient data distribution is finally done in section 4.3.

4.1 Performance Analysis of Optimal path Selection

Firstly, the proposed WF-CSO algorithm's performance is contrasted by the system with other conventional algorithms, namely, Chicken Swarm Optimization (CSO), Black Widow Optimization (BWO), Crow Search Algorithm (CSA), along with Genetic Algorithm (GA) concerning EED, network lifetime, PDR, and throughput. The simulated representation of every parameter is deliberated in the below table.

Table 2: Comparison of different methods

Metrics	Number of sensor nodes	Proposed WF-CSO	CSO	BWO	CSA	GA
End to End delay (ms)	20	0.21	0.6	0.9	1.1	1.34
	40	0.46	0.85	0.12	1.42	1.58
	60	0.69	0.98	0.154	1.66	1.72
	80	0.9	1.24	0.171	1.86	1.98
	100	1.2	1.58	0.2	2.1	2.4
Network lifetime (s)	20	510	351	210	150	105
	40	590	470	340	250	190
	60	730	550	460	390	320
	80	802	700	510	430	410
	100	910	730	620	520	480
Packet delivery ratio (%)	20	91	87	76	74	68
	40	89	84	71	69	64
	60	85	81	68	65	60
	80	81	77	64	61	58
	100	78	72	61	59	51
Throughput (kbps)	20	241	211	187	142	113
	40	234	200	172	136	101
	60	222	191	165	127	97
	80	212	180	157	121	83
	100	201	175	149	115	70

Discussion: The preceding table describes the proposed WF-CSO algorithm's performance, which is contrasted with the conventional CSO, BWO, CSA, and GA algorithms. Some performance metrics, say EED, network lifetime, PDR, and throughput metrics are utilized for doing the performance comparison. The performances are grounded upon several SN ranging from 20 to 100. The existing GA attained 1.34 ms EED, 105s network lifetime, 68 % PDR, and 113kbps throughput for '20' SN that offers lower performance than the proposed one. Lower performance is also offered by the existing CSO, BWO, and CSA. But, the proposed WF-CSO takes low EED (0.21ms), high network lifetime (510s), high PDR (91 %), and high throughput (241 kbps). Likewise, higher performance is offered by the proposed one when weighted against the existent methods aimed at the

remaining SN (40-100). Therefore, the system had given a better performance when weighted against the prevailing techniques as concluded by the discussion.

4.2 Performance Analysis of Congestion Detection

The suggested LM-NN's performance is contrasted with the traditional NN, ENN, SVM, and NB classifiers in this performance analysis of the congestion detection. Various parameters, namely recall, precision, accuracy, along with specificity metrics are deemed for assessing the suggested model's performance. This analysis could be graphically represented from the below figure,

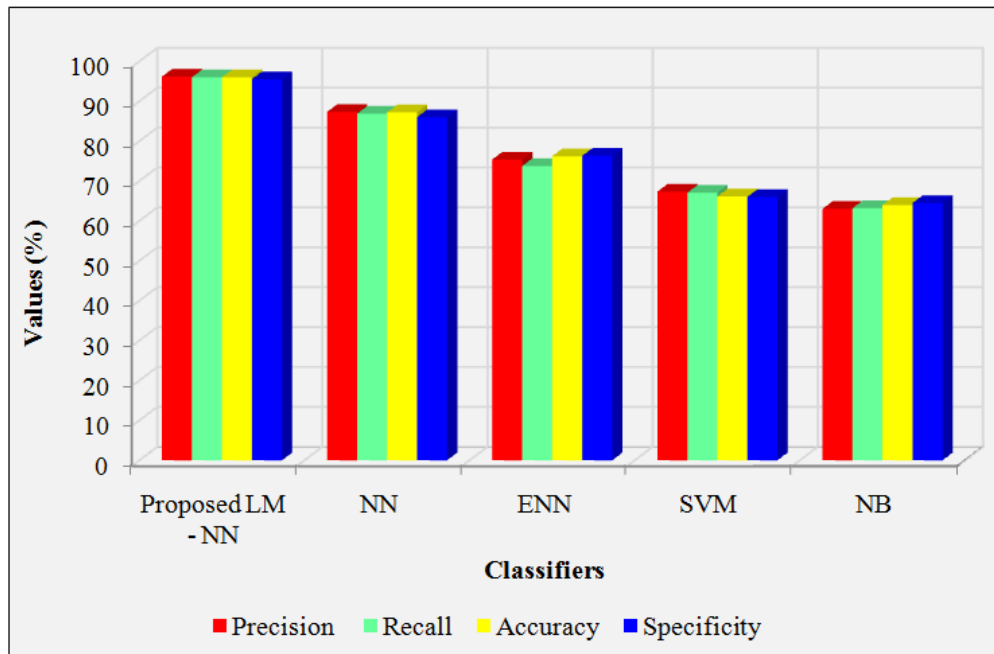


Figure 4: Comparative analysis of the proposed model with the conventional methods

Discussion: Figure 4 demonstrated the recommended LM-NN's performance with the traditional Neural Network (NN), Elman NN (ENN), Support Vectors Machine (SVM), along with Naive Bayes (NB) classifiers concerning the accuracy, recall, precision, along with specificity metrics. 63 % precision, 63.2 % recall, 64 % detection accuracy, and 64.5 % specificity are offered by conventional NB, which is very lesser when contrasted to the suggested one. A precision of 87.3 %, 75.38 %, and 67.32 %, recall of 86.9 %, 73.76 %, and 67.1 %, 87.2 %, 76.22 %, and 66.14 % detection accuracy, together with a specificity of 86 %, 76.32 %, and 66 % is offered by the conventional CSO, BWO, and CSA algorithm. A very high-performance value (i.e. 96.15 % precision, 96.01 % recall, 96 % accuracy, and 95.5 % specificity) is attained by the suggested one. Generally, the graph shows better performance results concerning precision, recall, accuracy, along with specificity utilizing the proposed LM-NN classifier.

4.3 Performance Analysis of Energy Efficient data Distribution

In this section, the proposed BD-SBO algorithm's performance is contrasted to the prevailing Satin Bowerbird Optimizer (SBO), Blue Monkey Optimization (BMO), Grasshopper Optimization Algorithm (GOA), and Particle Swarm Optimization (PSO) algorithms concerning Average Waiting Time (AWT), EC, Average Turn-around Time (ATT), along with CPU memory usage metrics. The performance is grounded upon the number of sensor data from 100 to 500. These are graphically explained from the below figure.

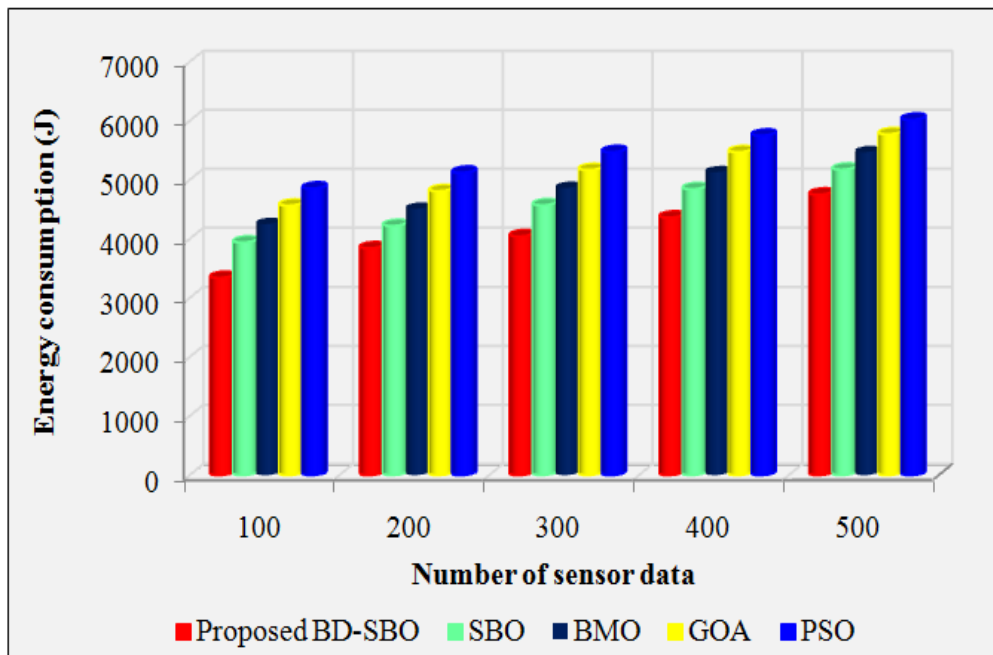


Figure 5: EC comparison graph

Discussion: The above figure indicates the proposed BD-SBO algorithm's performance with the conventional SBO, BMO, GOA, and PSO algorithms concerning EC metrics. The EC (average) is the main parameter for assessing the WSN's performance. It shows the EC throughout data distribution and the life span of the entire sensor network is predicted. For '100' nodes, the EC of 3342.25J is given by means of the proposed BD-SBO, which is lesser than the existent algorithms. Similarly, the proposed one offers the smallest EC values for the remaining sensor data (200-500). For an augmenting number of sensor data, the comparison analysis graph in Figure 5 reveals that the projected BD-SBO needs less energy when contrasted to the conventional SBO, BMO, GOA, and PSO algorithms.

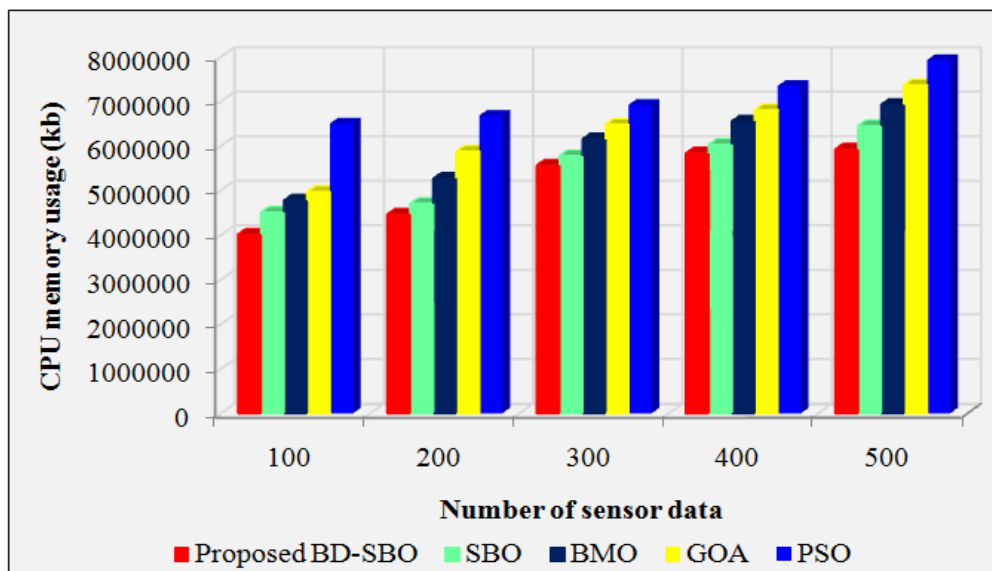
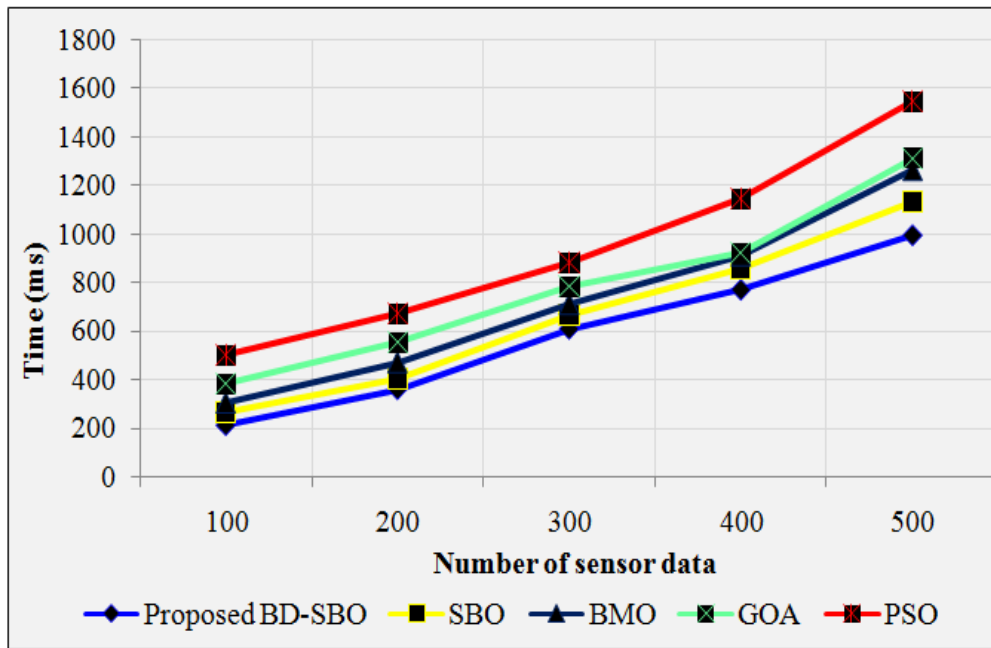


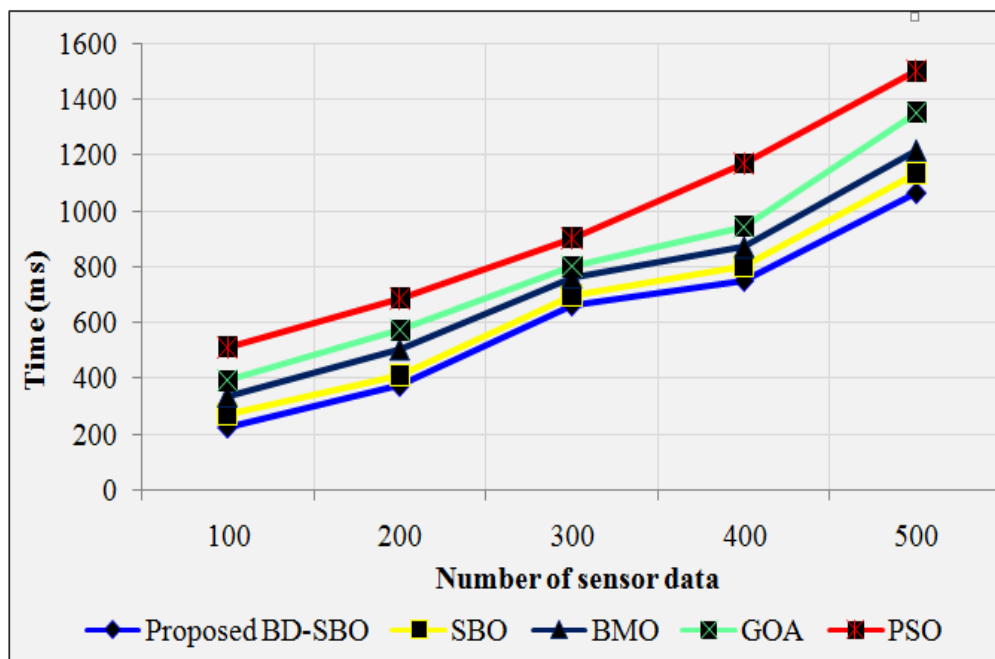
Figure 6: CPU memory usage analysis

Discussion: The above figure shows the proposed BD-SBO's performance, which is contrasted with the traditional SBO, BMO, GOA, and PSO algorithm grounded on CPU memory usage metrics. The proposed BD-SBO utilizes low CPU memory on deeming the prevailing approaches for 100 to 500 sensor data. The memory usage of the proposed work is increasing with the augmenting number of sensor data; but, the increased level does not exceed the existing techniques. For example, for 100-500 sensor data, the existent SBO, BMO, GOA, and PSO takes CPU memory of 6505960 kb, 6683224 kb, 6925408 kb, 7351000 kb, and 7930104 kb, but the proposed one takes 4015786 kb, 4465844 kb, 5563524 kb, 5835874 kb, and 5924578 kb, which is very lesser

when contrasted to the conventional techniques. The proposed BD-SBO's efficiency is proved from this analysis.



(a)



(b)

Figure 7: Comparative analysis of the proposed model with the existent methodologies in terms of (a) ATT and (b) AWT

Discussion: Figure 7 shows the resultant graph while contrasting the (a) ATT and (b) AWT with the conventional SBO, BMO, GOA, and PSO methods. From figure 7 (a), for 500 sensor data, an ATT of 1547ms is taken by the existing PSO, which is greater than the proposed one (since an ATT of 998ms is taken by the proposed one for '500' sensor data). From figure 7 (b), an AWT of just 1067ms is taken by the proposed one for 500 sensor data, but the existent PSO, GOA, BMO, and SBO takes an AWT of 1502ms, 1354ms, 1219ms, and 1135ms time respectively, which is greater when contrasted to the proposed one. Likewise, the proposed one consumes low ATT and AWT than all the existing methodologies for the remaining number of sensor data (100-

400). Therefore, the proposed one offers a higher-level performance than the existing techniques concerning ATT along with AWT as shown by the discussion.

5. CONCLUSION

The data communication performance between WSN towards cloud is affected immensely by the network congestion and sensor nodes' higher EC. This work proffered an energy-efficient data distribution process on the cloud utilizing the optimal routing path centred congestion control technique. The proposed system utilizes the LM-NN and WFCSSO aimed at detecting the congestion nodes, various routing paths, and detecting an optimal routing path aimed at energy-efficient routing. For offering energy-efficient data distribution on cloud data-centres, a BD-SBO is adapted. The proposed techniques' performance is analogized with numerous existent techniques for verifying the proposed system's effectiveness. The analogy outcome defines that utilizing the proposed algorithms the system acquires higher throughput, higher packet delivery ratio, higher network lifetime, and also less end to end delay. The WFCSSO decrements the end to end delay to 1.2ms aimed at 100 nodes and increments the network lifetime to 910 seconds. Alternatively, the proposed BD-SBO method employs only 4752.25J energy together with 5924578kb CPU memory to distribute 500 sensor data. The proposed LM-NN classifier precisely identifies the congestion nodes with a 96% accuracy-level. Hence, the proposed routing protocol is found to be highly accurate and also more energy-efficient. This work can be prolonged in the upcoming future by adding in cryptographic methods aimed at evading the attacks in data transmission.

REFERENCES

1. Pampapathi, B.M., Nageswara Gupta, M. & Hema, M.S. Data distribution and secure data transmission using IANFIS and MECC in IoT. *J Ambient Intell Human Comput* (2021), <https://doi.org/10.1007/s12652-020-02792-4>.
2. Mohammed Farsi, Mahmoud Badawy, Mona Moustafa, Hesham Arafat Ali, and YousryAbdulazeem, "A congestion-aware clustering and routing (CCR) protocol for mitigating congestion in WSN", *IEEE Access*, vol. 7, pp. 105402-105419, 2019, 10.1109/ACCESS.2019.2932951.
3. Aasem Ahmad, and ZdeněkHanzálek, "An energy efficient schedule for IEEE 802.15. 4/ZigBee cluster tree WSN with multiple collision domains and period crossing constraint", *IEEE Transactions on Industrial Informatics*, vol. 14, no. 1, pp. 12-23, 2017
4. KalpnaGuleria, and Anil Kumar Verma, "Comprehensive review for energy efficient hierarchical routing protocols on wireless sensor networks", *Wireless Networks*, vol. 25, no. 3, pp. 1159-1183, 2019.
5. JunaidAnees, Hao-Chun Zhang, BachirouGueneLougou, Sobia Baig, and YabibalGetahun Dessie, "Delay aware energy-efficient opportunistic node selection in restricted routing", *Computer Networks*, vol. 181, 2020, 107536. 10.1016/j.comnet.2020.107536.
6. Rong Du, LazarosGkatzikis, Carlo Fischione, and Ming Xiao, "On maximizing sensor network lifetime by energy balancing", *IEEE Transactions on Control of Network Systems*, vol. 5, no. 3, pp.1206-1218, 2017.
7. AmrutaLipare, Damodar Reddy Edla, and VenkatanareshbabuKuppili, "Energy efficient load balancing approach for avoiding energy hole problem in WSN using grey wolf optimizer with novel fitness function", *Applied Soft Computing*, vol. 84, pp. 105706, 2019, 10.1016/j.asoc.2019.105706.
8. Anxi Wang, Jian Shen, PandiVijayakumar, Yongxin Zhu, and Li Tian, "Secure big data communication for energy efficient intra-cluster in WSNs", *Information Sciences*, vol. 505, pp. 586-599, 2019, 10.1016/j.ins.2019.07.085
9. Nageswara Rao Maliseti, and Vinay Kumar Pamula, "Performance of Quasi Oppositional Butterfly Optimization Algorithm for Cluster Head Selection in WSNs", *Procedia Computer Science*, vol. 171, pp. 1953-1960, 2020, 10.1016/j.procs.2020.04.209
10. Kumaresan K and S. NithyaKalyani, "Energy efficient cluster based multilevel hierarchical routing for multi-hop wireless sensor network", *Journal of Ambient Intelligence and Humanized Computing*, 2020, 1-10. 10.1007/s12652-020-01700-0.
11. PrachiMaheshwari, Ajay K. Sharma, and Karan Verma, "Energy efficient cluster based routing protocol for WSN using butterfly optimization algorithm and ant colony optimization", *Ad Hoc Networks*, vol.110, pp.102317, 2020, 10.1016/j.adhoc.2020.102317.

12. Jiawei Tan, Wei Liu, Tian Wang, Shaobo Zhang, Anfeng Liu, Mande Xie, Ming Ma, and Ming Zhao, "An efficient information maximization based adaptive congestion control scheme in wireless sensor network", IEEE access, vol. 7, pp. 64878-64896, 2019, 10.1109/ACCESS.2019.2915385
13. Prasenjit Chanak, and Indrajit Banerjee, "Congestion Free Routing Mechanism for IoT-Enabled Wireless Sensor Networks for Smart Healthcare Applications", IEEE Transactions on Consumer Electronics, vol. 66, no. 3, pp. 223-232, 2020.
14. Wenguang Chen, Yugang Niu, and Yuanyuan Zou, "Congestion control and energy-balanced scheme based on the hierarchy for WSNs", IET wireless sensor systems, vol. 7, no. 1, pp. 1-8, 2016.
15. Nousheen Akhtar, Muazzam A. Khan Khattak, Ata Ullah, and Muhammad Younus Javed, "Efficient Routing Strategy for Congestion Avoidance in MANETs", In IEEE International Conference on Frontiers of Information Technology (FIT), pp. 305-309, 2017, 10.1109/FIT.2017.00061.
16. Yongjun Sun, Wenxin Dong, and Yahuan Chen, "An improved routing algorithm based on ant colony optimization in wireless sensor networks", IEEE communications Letters, vol. 21, no. 6, pp. 1317-1320, 2017.
17. Chuan Xu, Zhengying Xiong, Guofeng Zhao, and Shui Yu, "An energy-efficient region source routing protocol for lifetime maximization in WSN", IEEE Access, vol. 7, pp. 135277-135289, 2019, 10.1109/ACCESS.2019.2942321
18. Banavath Balaji Naik, Dhananjay Singh, and Arun Barun Samaddar, "Multi-objective Virtual Machine Selection in Cloud Data Centers Using Optimized Scheduling", Wireless Personal Communications, pp. 1-24, 2020, 10.1007/s11277-020-07807-z.
19. Anand Nayyar, and Rajeshwar Singh, "IEEMARP-a novel energy efficient multipath routing protocol based on ant Colony optimization (ACO) for dynamic sensor networks", Multimedia Tools and Applications, pp. 1-32, 2019, 10.1007/s11042-019-7627-z.
20. Loganathan D., M. Balasubramani, and R. Sabitha, "Energy Aware Efficient Data Aggregation (EAEDAR) with Re-scheduling Mechanism Using Clustering Techniques in Wireless Sensor Networks", Wireless Personal Communications. pp. 1-17, 2020, 10.1007/s11277-020-07985-w
21. Morteza Kheirkhah, Ian Wakeman, and George Parisi, "Multipath transport and packet spraying for efficient data delivery in data centres", Computer Networks, vol. 162, pp. 106852, 2019, 10.1016/j.comnet.2019.07.008
22. Amir Abbas Baradaran, and Keivan Navi, "HQCA-WSN: High-quality clustering algorithm and optimal cluster head selection using fuzzy logic in wireless sensor networks", Fuzzy Sets and Systems, vol. 389, pp. 114-144, 2020, 10.1016/j.fss.2019.11.015
23. Shishupal Kumar, Nidhi Lal, and Vijay Kumar Chaurasiya, "An energy efficient IPv6 packet delivery scheme for industrial IoT over G. 9959 protocol based wireless sensor network (WSN)", Computer Networks, vol. 154, pp. 79-87, 2019, 10.1016/j.comnet.2019.03.001.
24. Maurya S., Jain, V.K. and Chowdhury, D.R., "Delay aware energy efficient reliable routing for data transmission in heterogeneous mobile sink wireless sensor network", Journal of Network and Computer Applications, vol. 144, pp. 118-137, 2019, 10.1016/j.jnca.2019.06.012.
25. Deepak Mehta, and Sharad Saxena, "MCH-EOR: Multi-objective cluster head based energy-aware optimized routing algorithm in wireless sensor networks", Sustainable Computing: Informatics and Systems, vol. 28, pp. 100406, 2020, 10.1016/j.suscom.2020.100406.
26. Amit Singh and A. Nagaraju, "Low latency and energy efficient routing-aware network coding-based data transmission in multi-hop and multi-sink WSN", Ad Hoc Networks, vol. 107, pp. 102182, 2020, 10.1016/j.adhoc.2020.102182.
27. Pampapathi B.M., M. Nageswara Guptha, and Mahantesh H M, "Survey on Data Communication Frameworks in IoT." International Journal of Management, Technology and Engineering (IJMTE) Volume IX, Issue VI, ISSN NO: 2249-7455 in June 2019.
28. Pampapathi B.M., M. Nageswara Guptha, and M.S. Hema, "Malicious Node Detection and Energy-aware Optimal Routing in Wireless Sensor Networks using CD-LVQ and BMSSO Algorithms". The Journal of Huazhong University of Science and Technology (2021): <http://hustjournal.com/vol50mar-2>.
29. Mahantesh, H.M., Nageswara Guptha, M. & Hema, M.S. Optimized Path and Reduced Rule Caching Cost for Software Defined Network (SDN) Based Internet of Things (IOT). Wireless Pers Commun (2021). <https://doi.org/10.1007/s11277-021-08698-4>.