International Journal of Reconfigurable and Embedded Systems (IJRES)
Vol. 15, No. 3, November 2026, pp. 848~858
ISSN: 2089-4864, DOI: 10.11591/ijres.v15.i3.pp848-858

Energy efficient data gathering in load balanced clustered wireless sensor networks with compressive sensing

Raju Chitra[1], Rajagopal Sudarmani[1], Kannadhasan Suriyan[2]
[1]Department of Electronics and Communication Engineering, School of Engineering, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, India
[2]Department of Electronics and Communication Engineering, Study World College of Engineering, Coimbatore, India

Article Info

Article history:

Received Jun 26, 2025
Revised Jun 17, 2026
Accepted Jul 18, 2026

Keywords:

Bernoulli holomorphic
Compressive sensing
Data gathering
Hermite polynomial
Swarm optimization
Wireless sensor network

ABSTRACT

Rapid energy loss from continuous sensing and data transmission often shortens network lifespan. This work suggests Bernoulli holomorphic swarm optimized Hermite polynomial compressive sensing (BHSO-HPCS), an energy-efficient data collection method, to overcome this difficulty. The suggested approach minimizes energy usage without sacrificing network performance by combining compressive sensing (CS) and load-balanced clustering algorithms. Sensor nodes (SNs) initially support large-scale wireless sensor network (WSN) deployments. Cluster heads (CHs) are chosen from nodes with relatively more remaining energy. The Bernoulli holomorphic (BH) balanced iterative reducing and clustering using hierarchies (BIRCH)-Tchebychev particle swarm optimization model is then used to create load-balanced clusters. A BH radius-of-convergence method improves the standard cluster threshold, and the Tchebychev function minimizes reaction time during data aggregation by optimizing average communication distance. In order to reduce communication overhead, sensor data is sent to CHs before being sent to the mobile sink. HPCS is also used for effective redundancy reduction and data aggregation. This extends the network's lifespan and boosts throughput. Simulation results using network simulator 2 (NS2) show that BHSO-HPCS significantly outperforms circular clustering using the hybrid crow search algorithm (HCSA) and weighted data aggregation trees with optimum mobile sink(s) (WDAT-OMS) in terms of energy consumption, reaction time, throughput, and overall network efficiency.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Raju Chitra
Department of Electronics and Communication Engineering, School of Engineering
Avinashilingam Institute for Home Science and Higher Education for Women
Coimbatore, Tamil Nadu, India
Email: chitra_ece@avinuty.ac.in

1. INTRODUCTION

Wireless sensor networks (WSNs) have emerged as a dynamic worldwide characteristic in recent decades. Their applications have been extended into a wide range of fields, including forest fire detection, video surveillance, healthcare data analytics, and landslide monitoring at vulnerable places, thanks to wireless communications. The most crucial WSN task is data collecting. One of the most crucial criteria for WSN is still energy-efficient data collecting. This is due to the idea that gathering data would reduce energy use. The multiple sensor nodes (SNs) that make up WSN are in charge of collecting data and sending it to the base station (BS). One of the challenges with WSNs is the energy limitation, which negatively impacts the

network's lifespan. Numerous energy-efficient methods have been developed recently. An energy-efficient data gathering method using circular clustering and the hybrid crow search algorithm (HCSA) was described in order to extend the network's lifetime [1].

The whole sensor network was initially divided in order to produce the circular cell clusters in the proposed technique. In order to enable precise data collecting in the circular cell cluster region based on proximity, a multi-objective weighted sum model was also created. Data was also gathered from the cluster heads (CHs) via HCSA, routing, and a dynamic mobile sink relocation mechanism. Both total energy consumption and network longevity were decreased as a result of these techniques. However, there was no aim for the data gathering response time. Weighted data aggregation trees with optimal mobile sink(s) (WDAT-OMS) is a data aggregation approach that was established to reduce energy consumption and latency in WSN [2]–[5]. The creation of the WDAT-OMS approach was based on a two-level design. At the first level, sensors were gathered together. In addition to load-balanced data aggregation trees, packets from sensors were routed to the proper CHs at the second level using compression sensing theory. Finally, to retrieve the aggregated data, one or more sinks traverse the network, boosting scalability and balancing energy consumption. Despite advances in energy consumption and scalability, the throughput was not investigated. For heterogeneous clustered-based WSNs, a secure and load-balanced routing (SLBR) method was put forward [6]–[11].

Longer delays and energy usage restrictions still need to be addressed, despite the research community's optimistic attempts to find continuous node mobility. Given these problems, publications [12]–[16] focused on a hybrid data collection approach that included both node locations and network characteristics. With this kind of network design, it was claimed that load balancing would lead to reduced energy use. High expenses have resulted from the development of many innovative methods to enhance balanced energy via the use of devices equipped with energy harvesters. These variables need the development of a practical plan for balancing energy use. To address this problem, a wireless energy balancer-based energy-efficient clustering technique was developed [17]–[20]. This kind of architecture resulted in a significant decrease in network overhead and computation time. A novel software-defined networking (SDN) architecture was developed in order to decrease the load distribution across SNs and hence extend the network lifetime. In this instance, additional variables including topology, virtual routing, and connections were considered. A novel load-balanced routing method that balanced network consumption and greatly increased network longevity was proposed as a result of this discovery. A data aggregation method based on compressive sensing (CS) has been shown to be successful in improving data aggregation performance and network lifetime. The temporal connection based on the physical features of the SN's location is included into each subsequent observation. These kinds of spatiotemporal connections, according to researchers, effectively improve energy savings. A spatial and temporal correlation-based data redundancy reduction (STCDRR) technique was proposed to eliminate redundancy at the source and aggregator levels. Another scalable and load-balanced method for data aggregation employing mobile agents for WSN was proposed. This kind of design improved network longevity and energy consumption while decreasing response time. However, the duplication that comes with data aggregation received little attention. To address this problem, elliptic curve cryptography with CS was proposed [21]–[25]. Here, energy consumption, network lifespan, and security were all reported to have improved.

2. MATERIALS AND METHODS

This work proposes Bernoulli holomorphic swarm optimized Hermite polynomial compressive sensing (BHSO-HPCS). The suggested BHSO-HPCS approach, which attempts to resolve the concerns brought up in the preceding parts, is described in this section. The working area is divided into three stages by the BHSO-HPCS: data gathering, cluster construction, and SN deployment.

The detailed description is provided in the following sections. Let us consider a WSN comprising of 'n' SNs '$(S_1, S_2, \dots, S_n)$' in a two-dimensional network field with radius, 'r' and also presume that all SNs communicate with the 'CH' only after initial SN deployment possessing same sensing range 'SR' respectively. First, SN deployment is modelled at random. We then present the received signal strength indication (RSSI) radio model that effectively captures distinguishing characteristics (i.e., transmission power, path loss, and fading) and provide empirical portrayal between SN communications in random deployment to ensure optimal SN deployment. Link quality is measured by signal strength using the RSSI radio model. The RSSI radio model determines node adjacency during communication by measuring the separation between two SNs. A wireless signal's intensity is represented by the RSSI radio model, which is described as (1).

$$RSSI = -10 * PL * log_{10}(Dis) * T \quad (1)$$

From (1), the link quality '$RSSI$' of each SN is measured based on the transmission power 'T', distance between SNs 'Dis' and path loss 'PL', respectively. Next, the distance based on polar coordinates is measured to improve coverage aspect after random deployment of SN. During initialization of SN deployment, the sink node 'SN' measures polar coordinate position and broadcasts that information to every SN in the WSN. Let us consider the polar coordinates of 'x' be '(a, θ)' and polar coordinates of 'y' be '(b, φ)', then the distance is measured as (2).

$$Dis(x, y) = \sqrt{a^2 + b^2 - 2ab(\theta - \varphi)} \quad (2)$$

From (2), the distance 'Dis' between SNs 'x, y' is measured based on the law of cosines (i.e., length of the side of triangle 'φ' to the cosine of one of its angles 'θ', respectively). By means of this distance measurement, the overlap in node sensing area is reduced and therefore results in the improvement of overall network coverage. In WSN, every SN possesses similar energy during node deployment. Then the SN consumes certain amount of energy for sensing, processing, and data routing. The initial energy of SN is measured on the basis of power and time and is formulated as (3).

$$E_n = P[S_i] * \mathrm{T}[S_i] \quad (3)$$

From (3), the initial energy 'E_n', is obtained employing power '$P[S_i]$' and the time 'T $[S_i]$', respectively. In addition, the residual energy for each SN is also measured and is mathematically represented as (4).

$$E_{res} = E_{ini} - E_{cons} \quad (4)$$

From (4), the residual energy 'E_{res}', of SN is measured taking into consideration initial energy of SN 'E_{ini}' and consumed energy of SN 'E_{cons}', respectively. The SN that has more energy than other SNs is thus referred to as the CH node. The SN deployment is shown in Figure 1. As can be seen from the above, Figure 1(b) displays the final sensor placements as established by the RSSI radio model, while Figure 1(a) displays the beginning sensor positions based on the random deployment model. The numbers in the circle centers represent the identification (i.e., SN IDs) of the ninety-nine SNs that were placed in the sensor field. Ten CHs were chosen based on energy-efficient SNs and the RSSI radio model. The next stage is to build a cluster in a load-balanced manner with an energy-efficient CH chosen. Here, the clustering feature (CF) tree is built using the Bernoulli holomorphic (BH) function, with the BH balanced iterative reducing and clustering using hierarchies (BIRCH) model divided into two phases. Tchebychev distance and load-based fitness are then used to create global clustering (GC), guaranteeing load-balanced cluster formation. To start with, initially, each particle (i.e., SN) 'S_i' is allocated with a position 'Pos_{id}' and velocity 'Vel_{id}' in an arbitrary fashion and '$d = (1, 2, \dots, D)$', respectively. During each iteration's, each particle or SN fine tunes its velocity based on two best solutions. The first solution being the best solution identified so far, referred to as 'lb_i' and the second solution being the global best 'gb', respectively. Upon successful identification of the two best values, i.e., local best and global best, the SN 'S_i' update its position and velocity as (5) and (6).

$$Vel_{id}[S_i] = W * Vel_{id}(t) + c_1 * a_1 * [lb_i(t) - Pos_{id}(t)] + c_2 * a_2[gb(t) - Pos_{id}(t)] \quad (5)$$

$$Pos_{id}(t + 1)[S_i] = Pos_{id}(t) + Vel_{id}(t + 1) \quad (6)$$

From (5) and (6), the velocity 'Vel_{id}' and position 'Pos_{id}' of each SN 'S_i' is updated based on the learning factors 'c_1', 'c_2' and the arbitrary values 'a_1', 'a_2', with weight 'W' controlling the SN velocity. The BH BIRCH model is performed in two different phases, building CF tree and GC. The BH BIRCH model uses a tree structure to create a cluster. In the CF tree, CF is denoted as given.

$$CF \rightarrow (S_i, VS[FD(S_i)], [FD(S_i)^2]) \quad (7)$$

From (7), the clustering feature 'CF', is denoted based on the number of data samples (i.e., SNs) 'S_i' in the CF, vector sum of feature dimensions of data samples (i.e., SNs) '$VS[FD(S_i)]$' in the CF and the square of feature dimensions of data samples (i.e., SNs) '$FD(S_i)^2$', respectively. The CF tree includes an essential criterion, threshold. In our work, the conventional form of threshold representing maximum number of data samples (i.e., SNs) in a sub-cluster of the leaf node is modified by employing BH radius of convergence, where instead of the diameter of the cluster, radius from central point is employed for acquiring the threshold. To determine the maximum number of data samples or SNs in a sub-cluster of the leaf node, the nearest point in the complex plane by employing radius of convergence is obtained as (8).

$$f(S) = \frac{1}{1+S^2} \quad (8)$$

With the above obtained nearest point in the complex plane for each SN 'S' as given in (8), the complex valued function 'f' of a single complex SN, the derivative of 'f' at a point 'S_0' using BH function is mathematically represented as given.

$$f'(S_i) = \lim_{S \to S_i} \frac{f(S_i) - f(S_j)}{S_i - S_j} \quad (9)$$

With the above results from (9), maximum number of data samples or SNs in a sub-cluster of the leaf node is arrived at in an efficient manner. Next, in the second phase, GC is performed i.e., load balanced cluster formation is modeled by utilizing Tchebychev particle swarm optimization (PSO). Here, in addition to the basic formulation of PSO as given in (5) and (6), fitness estimation is made by optimizing average distance using Tchebychev distance function and average energy of SNs from CH and overall SNs under consideration. Then, the Tchebychev distance between two SNs with coordinates 'S_i', 'S_j' is mathematically represented as given.

$$Dis_{Tchebychev}(S_i, S_j) = Max(|S_i - S_j|) \quad (10)$$

From (10), Tchebychev distance between two SNs is arrived at based on the distance between two vector coordinates with the largest of their variances '$Max(|S_i - S_j|)$' along any coordinate dimension.

$$fitness = Expected\ CH * \frac{Tot(CH)}{Dis_{Tchebychev}(S_i, S_j)} \quad (11)$$

It is clear from (11), that the proposed fitness function considers distance between two vectors coordinates with the largest of their variances in addition to the expected CH factor and the total number of CHs to judge the quality of the sensors. SNs transmit data to the CHs in a clustered WSN. The amount of SNs in the queue determines how many SNs are allotted. This helps to maximize network lifespan and minimize energy use. Therefore, the suggested fitness function that uses Tchebychev distance in GC is beneficial for both energy efficiency and load balancing. Lastly, by sending each SN's data packet to its CH rather than directly passing it to the mobile sink, the load balanced cluster creation with cluster feature building utilizing BH BIRCH and GC via Tchebychev distance reduces the reaction time throughout the data aggregation process. The following is Algorithm 1 for BH BIRCH-Tchebychev PSO-based cluster formation:

Algorithm 1. BH BIRCH-Tchebychev PSO-based cluster formation
Input: Sensor node '$S = \{S_1, S_2, \dots, S_n\}$'
Output: Energy-efficient Load-balanced Cluster Formation
Step 1: **Initialize** 'n', transmission power 'T', path loss 'PL', '$d = (1,2, \dots, D)$',
Step 2: **Initialize** polar coordinates of 'x' be '(a, θ)' and polar coordinates of 'y' be '(b, φ)'
Step 3: **Initialize** '$c_1, c_2 = 2$', '$a_1 a_2 = [0,1]$'
Step 4: **Initialize** length of the side of triangle 'φ' to the cosine of one of its angles 'θ' respectively
Step 5: **Begin**
Step 6: **For** each Sensor node 'S'
Step 7: Measure RSSI radio based on distance using (1) and (2)
Step 8: Measure initial energy of SN using (3)
Step 9: Measure residual energy using (4)
Step 10: **Return** higher residual energy as cluster head 'CH'
Step 11: **End for**
Step 12: **For** each Sensor node 'S' with cluster head 'CH'
Step 13: Update position and velocity of each sensor node based on local best and global best using (5) and (6)
//Clustering Featuring
Step 14: Formulate clustering feature using (7)
Step 15: Determine maximum number of data samples or SNs in a sub-cluster of the leaf node using (8) and (9)
//Global clustering
Step 16: Evaluate Tchebychev distance between two SNs using (10)
Step 17: **Return** load balanced cluster 'Cl'

Step 18: **End for**
Step 19: **End**

According to the aforesaid method, CHs are chosen based on energy-efficient features and the RSSI radio model. The local best and global best solutions are then used to update each SN's location and velocity. The BH maximum number of data samples or SNs in a sub-cluster of the leaf node is used in the third cluster feature to calculate the maximum number of SNs in a sub-cluster of the lead node. Lastly, Tchebychev distance is used to design GC, guaranteeing load-balanced cluster formation. The response time is reported to be decreased in this way by sending each SN's data packets to its CH rather than directly forwarding to the mobile sink. Lastly, the fitness function is used to improve the routing of sensed data packets, and the dynamic source routing (DSR) routing protocol is used to transmit sensed data to CH. The Hermite polynomial CS is used as a data aggregation model, and its projections—which are then stated to be regenerated at a later stage—maintain the data packets contained in the CH. The performance or throughput of the aggregation process and the overall network lifetime with minimal redundancy are improved by using the Hermite polynomial function via CS, which preserves the important data packets in the original data packets and decreases the components of redundancy (i.e., same SN data packets at the CH).

Let us assume that the input variable (i.e., the load balanced cluster) '$Cl = [Cl_1, Cl_2, \dots, Cl_d]$', the sensed data of all SNs, is a '$d - dimensional\ variable$' with all data packets following independent normal distribution. Then, the Hermite orthogonal polynomial is mathematically formulated as (12) the structure of Hermite polynomial CS-based data aggregation model.

$$X = \sum_{I \le p} O_I \alpha_I(Cl) \tag{12}$$

From (12), 'I' refers to the multi-layered index with orthogonal order of polynomial being 'p' and multivariable (i.e., data packets obtained from different sensors, temperature, and humidity) orthogonal Hermite function 'α_I' as (13).

$$\alpha_I(Cl) = \left\{h_{i_1}(Cl_1), h_{i_2}(Cl_2), \dots, h_{i_d}(Cl_d)\right\} \tag{13}$$

From (13), '$h_i(.)$' represent the univalent Hermite function. The objective here remains in obtaining the orthogonal coefficient 'O_I' and the number of SNs with redundant data packets in 'O_I'. This is mathematically expressed as (14).

$$P = \frac{(p+d)!}{p!d!} \tag{14}$$

Finally, the Hermite orthogonal polynomial from (12) is rewritten as (15).

$$X = \sum_{i=1}^{P} O_I \alpha_I \tag{15}$$

After the expansion of 'X' is acquired, the orthogonal coefficient '$O = O_{ij}$' of 'X' is reconstructed by means of CS. Let us consider 'h' arbitrary clusters '$(Cl_1, Cl_2, \dots, Cl_h)$' of '$Cl$', to obtain corresponding sample instances '$s = (X_1, X_2, \dots, X_h)^T$' as (16).

$$s = \alpha O' \tag{16}$$

From (16), 'α' is referred to as the quantification or weighing matrix that is arrived at by putting forward different Hermite orthogonal polynomials from arbitrary sample instances '$(Cl_1, Cl_2, \dots, Cl_h)$' as (17) and (18).

$$\alpha = \begin{bmatrix} \alpha_1(Cl_1) & \alpha_2(Cl_1) & \alpha_3(Cl_1) & \dots & \alpha_p(Cl_1) \\ \alpha_1(Cl_2) & \alpha_2(Cl_2) & \alpha_3(Cl_2) & \dots & \alpha_p(Cl_2) \\ \dots & \dots & \dots & \dots & \dots \\ \alpha_1(Cl_h) & \alpha_2(Cl_h) & \alpha_3(Cl_h) & \dots & \alpha_p(Cl_h) \end{bmatrix} \tag{17}$$

$$O' = \left[O_1, O_2, \dots, O_p\right]^T \tag{18}$$

Supplementary restrictions require to be exploited with the purpose of ensuring the distinctiveness of the SNs data packets being gathered at the CH. This is achieved by minimizing the orthogonal coefficient in such a manner that the multivariable orthogonal Hermite function should be equal to the sample instances. This is mathematically expressed as (19).

$$\min(O'), such\ that\ \alpha O' = s \quad (19)$$

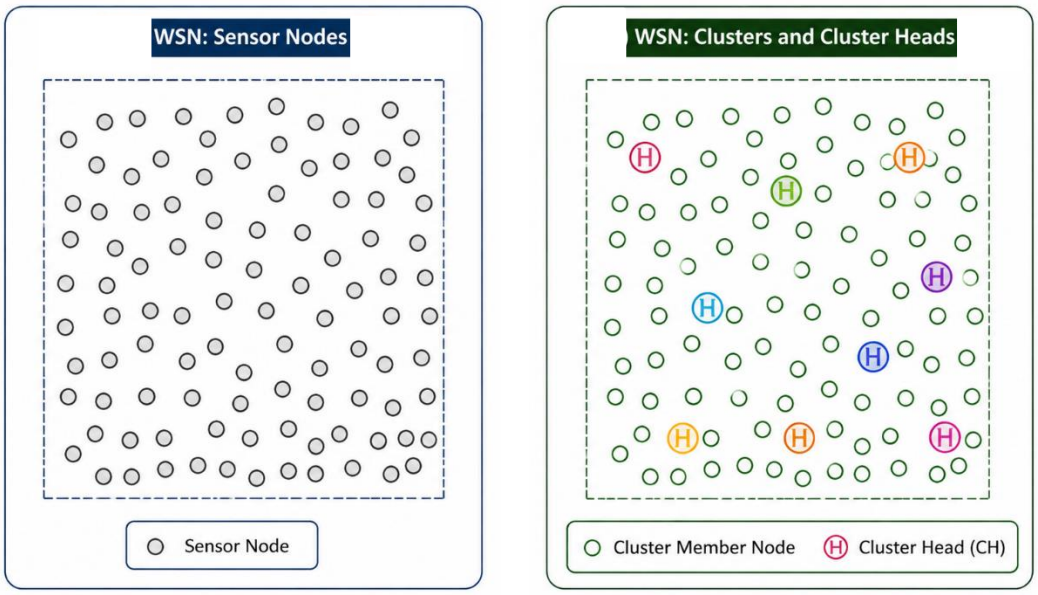
Each SN utilizes the Hermite functionality test to determine the difference between the previous period of data packets and the present data packets during the data (packet) aggregation phase at the CH. It just discards the new data packet and simply sends a disclosure to the CH after determining that the change is not substantial. The disclosure data packet is empty, indicating that the new and old data packets are equivalent in order to avoid transmitting duplicate data packets, which would save energy consumption and extend the lifespan of the network. The pseudocode representation of Hermite polynomial CS-based data aggregation is provided in accordance with the preceding technique.

The suggested HPCS-based data aggregation process is shown in Algorithm 2. Its goal is to minimize duplicated sensor data and the quantity of information carried across the network in order to lower communication overhead and extend network lifespan. First, load-balanced clusters with matching CHs are found after the SNs are arranged into random clusters (Steps 1-3). Equations are used to create the Hermite orthogonal polynomial for each SN connected to a CH (12)–(15) (Step 4). After evaluating the associated orthogonal Hermite functions for each load-balanced cluster using (13) and (14) (Step 5), a compact representation of the sensed information is obtained by reformulating the Hermite polynomial representation using (15) (Step 6). This transformation effectively represents correlated sensor values while minimizing needless data transfer by taking use of the orthogonal properties of Hermite polynomials. Compressive reconstruction and redundancy reduction are carried out in Algorithm 2's second stage (Steps 8–12). The (17) is used to build the weighting matrices for each random cluster that contains the chosen sample instances, and (18) is used to generate the necessary sample instances (Steps 8–9). In order to find and reduce duplicate information among the combined sensor readings, the multivariable orthogonal Hermite function derived in (19) is then assessed (Step 10). The reduced CS representation (Step 11), which includes the crucial sensing data with fewer measurements, is the final output of the process. In order to reduce data redundancy and transmission needs, Algorithm 2 combines cluster-based aggregation, Hermite polynomial transformation, compressive sampling, and reconstruction. As a consequence, packet transmission and processing overhead as a result, the network's overall lifespan improves, and SN energy consumption decreases due to reduced packet transmission and processing overhead, which lowers SN energy consumption and improves the lifespan of the network as a whole.

Algorithm 2. Hermite polynomial CS-based data aggregation
Input: Sensor node '$S = \{S_1, S_2, \dots, S_n\}$'
Output: reduced compressive sensing results with improved network lifetime
Step 1: **Initialize**'h' arbitrary clusters '$(Cl_1, Cl_2, \dots, Cl_h)$' of '$Cl$'
Step 2: **Begin**
Step 3: **For** each Sensor node 'S' with cluster head 'CH' and load balanced cluster 'Cl'
Step 4: Formulate Hermite orthogonal polynomial for each load balanced cluster as given in (12), (13), (14) and (15)
Step 5: Evaluate orthogonal Hermite function for each load balanced cluster as given in (13) and (14)
Step 6: Rewrite the Hermite orthogonal polynomial from as given in (15)
Step 7: **End for**
//Reconstruction
Step 8: **For** each arbitrary clusters '$(Cl_1, Cl_2, \dots, Cl_d)$' of '$Cl$'with sample instances '$s = (X_1, X_2, \dots, X_h)^T$
Step 9: Formulate weighing matrices as given in (17) and obtain arbitrary sample instances as given in (18)
Step 10: Evaluate multivariable orthogonal Hermite function function for minimizing redundancy as given in (19)
Step 11: **Return** reduced compressive sensing results
Step 12: **End for**
Step 13: **End**

The aforementioned approach applies CS via Hermite polynomial orthogonal function to the sample instances acquired as input in order to minimize redundancy for data packet aggregation in WSN. Reconstruction is then carried out in a way that minimizes duplication, with validation testing based on a disclosure given to CH. Figure 1 illustrates SN deployment and CH selection in a WSN. Figure 1(a) shows the random deployment of SNs, while Figure 1(b) illustrates formation of clusters and selection of CHs based

on RSSI for load-balanced data aggregationn. As a consequence, there are fewer duplicated data packets, which significantly reduces energy usage and network lifespan.



(a) (b)

Figure 1. SN deployment and CH selection in a WSN; (a) random deployment and (b) RSSI radio model

3. RESULTS AND DISCUSSION

The network simulator 2 (NS2), an Intel(R) Core (TM) i3 processor with a 3.40 GHz CPU and 4 GB RAM running Microsoft Windows 8, is used for the trials. The studies are conducted under the assumption of a WSN scenario with a number of SNs and gateways placed in a 200×200 Sqm region, each having a 100 m sensing range. Table 1 displays the WSN simulation settings.

Table 1. Simulation parameters

Parameter	Value
Number of SNs	50, 100, 150, 200, 250, 300, 350, 400, 450, and 500
Area	200×200 sqm
Communication range	100 m
Simulation time	200 seconds
Simulation runs	10
Transmission duration	150 seconds
Initial energy	3 Joules

The suggested BHSO-HPCS approach was compared to two cutting-edge data aggregation techniques: WDAT-OMS [2] and circular clustering with HCSA [1]. Similar amounts of SNs and data packets were used for simulation, and an average of ten runs were carried out to validate the efficiency in order to guaranty a fair comparison between the suggested BHSO-HPCS and current approaches [1], [2]. In terms of energy consumption, reaction time, throughput, and network lifespan, the suggested load balanced data aggregation method's performance is assessed. This total energy consumption determines how much energy the SNs need overall to move the data packets through the simulation. It is determined by calculating how much energy the SNs need to move data packets from the source to the mobile sink node. The SN's total energy usage is expressed as (20).

$$CS_E = \sum_{i=1}^{n} S_i * CS_E\ (SS) \quad (20)$$

According to (20), the energy consumption "〖CS〗_E" is calculated based on the number of SNs participating in the simulation process "S_i" and the actual energy consumption of individual SNs "〖CS〗

_E (SS)," respectively. Joules (J) are used to measure the total amount of energy used. Table 2 displays the energy usage findings from (20). Measuring the amount of energy used (in Joules) per round is regarded as an important performance parameter. Circular clustering, HCSA [1], WDAT-OMS [2], and BHSO-HPCS are used to compare different numbers of SNs (50, 100, and 150) with equal and unequal loads. Figure findings show that the suggested BHSO-HPCS approach outperforms [1], [2]. The total energy consumption was found to be 100 J, 150 J, and 175 J, respectively, in the simulations conducted with 50 SNs using the suggested BHSO-HPCS technique and current [1], [2]. As a consequence, it was found that the energy usage employing the BHSO-HPCS approach was superior than [1], [2]. The BH BIRCH-Tchebychev PSO model's load-balanced cluster formation nature is the cause.

It is due to the proposed fitness function employing Tchebychev distance that takes into consideration the average energy of SNs from CH and overall SNs under consideration while assigning the sensors. Hence, the solution produced by proposed BHSO-HPCS method is energy efficient by 31% compared to [1] and 44% compared to [2], respectively. It is measured as the ratio of the number of data packets transferred per unit time. In other words, throughput is measured on the basis of number of data packets received at the destination versus the total number of data packets sent from the source. The throughput is calculated as (21):

$$T = \frac{DP_{received}}{DP_{sent}} * 100 \quad (21)$$

From (21), the throughput success rate 'T' is measured based on the data packets received '$DP_{received}$' to the data packets sent 'DP_{sent}'. It is measured in terms of percentage (%). The results of throughput performed for an average of 10 simulation runs from (21) are listed in Table 3.

Table 2. Energy consumption comparison

SNs	Energy consumption (J) BHSO-HPCS	circular clustering and HCSA	WDAT-OMS
50	100	150	175
100	125	185	215
150	140	215	240
200	165	235	295
250	190	250	335
300	210	315	385
350	225	345	425
400	240	380	485
450	295	415	525
500	335	435	585

Table 3. Throughput comparison

SNs	Throughput (%) BHSO-HPCS	Circular clustering and HCSA	WDAT-OMS
50	92	88	84
100	90.35	87.25	83.15
150	90	86.85	83
200	89.15	86	82.85
250	89	85.45	82.55
300	88.45	85	81
350	87	84.25	80.35
400	85.25	83.15	79.25
450	83.15	80	78
500	81	79	77

The performance comparison findings in terms of throughput employing three distinct approaches—BHSO-HPCS, circular clustering, HCSA [1], and WDAT-OMS [2]—are shown in Table 3. As the number of SNs between source and destination increases, the network traffic throughput decreases for each size of data packet. This is because throughput decreases as the number of SNs grows due to an increase in the data packet error probability. However, the throughput rates for the three methods—BHSO-HPCS, [1], and [2]—were found to be 92%, 88%, and 84%, respectively, when simulations were run for 25 data packets using 50 SNs. The throughput rate employing the suggested BHSO-HPCS approach was determined to be somewhat superior than [1], [2] based on the simulation results. By eliminating data redundancy by the use of Hermite polynomial CS-based data aggregation, the throughput rate was improved by 4% and 8% in comparison to [1], [2], respectively. The time at which the first SN failure results from battery power charge discharge is measured by this network lifetime performance parameter. More precisely, the number of SN in the network after recurrent data aggregation in WSN is used to calculate the network lifespan. The network lifespan may be expressed numerically as (22).

$$NL = \left(\frac{SN_{addressed}}{n}\right) * 100 \quad (22)$$

From (22), the network lifetime 'NL' is measured using the total number of SNs 'n' in the network and the SN addressed '$S_{addressed}$' for data aggregation in WSN. Higher the network lifetime, more efficient the method is said to be. Here, the network lifetime is derived as percentage node alive (PNA). Finally, the results of network lifetime for an average of 10 simulation runs from (22) are listed in Table 4.

Table 4. Network lifetime comparison

SNs	Network lifetime (%)		
	BHSO-HPCS	Circular clustering and HCSA	WDAT-OMS
50	92	88	86
100	91.35	87.45	85.45
150	90	87	84
200	89.55	86.25	83.25
250	88.15	86	82
300	87.65	85.45	81.85
350	87.15	83	81
400	85.35	82.15	80.65
450	84	80.45	78.35
500	82.35	79.55	77

The entire number of iterations from the beginning of a WSN's operation until the final SN dies is known as the network lifespan. In our study, network lifespan is expressed as a percentage (%). Experiments are conducted utilizing various numbers of SNs ranging from 50 to 500, as seen in the above figure. We have conducted tests utilizing circular clustering, HCSA [1], WDAT-OMS [2], and the suggested BHSO-HPCS approach for comparison. The graphic makes it evident that the suggested BHSO-HPCS approach has a longer network life than [1], [2]. This is because the suggested reconstruction function takes into account both the CS-based data aggregation and the Hermite functionality test, which measures the difference between the previous and current periods of data packets. As a result, the suggested BHSO-HPCS method's total network lifespan is improved by 4% over [1] and 7% over [2].

4. CONCLUSION

In order to execute aggregated data at CH in WSN with minimal data redundancy and load balanced cluster creation, this study proposes BHSO-HPCS. The complete WSN is first divided according to the RSSI radio model in the suggested approach for guaranteeing an energy-efficient data aggregation node portion. The BH BIRCH-Tchebychev PSO model is then created to guaranty load-balanced cluster formation. It minimizes energy consumption as well as response time involved in data aggregation with the help of the BH function in the cluster featuring construction tree, Tchebychev distance, and load-based fitness in GC. Additionally, a data aggregation technique based on HPCS is used to decrease redundancy at CH. By creating a Hermite orthogonal polynomial and carrying out reconstruction with the least amount of redundancy, the fitness function is altered in this way, increasing throughput and network longevity. The testing findings shown that the suggested approach significantly outperformed in terms of energy usage, reaction time, throughput, and network longevity.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Raju Chitra	✓	✓	✓	✓		✓	✓	✓	✓	✓				
Rajagopal Sudarmani	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓		
Kannadhasan Suriyan	✓	✓	✓	✓		✓	✓		✓	✓	✓			

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

Data availability is not applicable to this paper as no new data were created or analyzed in this study.

REFERENCES

[1] P. K. Kodoth and G. Edachana, "An energy efficient data gathering scheme for wireless sensor networks using hybrid crow search algorithm," *IET Commun.*, vol. 15, no. 7, pp. 906–916, 2021, doi: 10.1049/cmu2.12128.

[2] S. P. Tirani, A. Avokh, and S. Azar, "WDAT-OMS: A two-level scheme for efficient data gathering in mobile-sink wireless sensor networks using compressive sensing theory," *IET Commun.*, vol. 14, no. 11, pp. 1826–1837, Jul. 2020, doi: 10.1049/iet-com.2019.0433.

[3] G. Thahniyath and M. Jayaprasad, "Secure and load balanced routing model for wireless sensor networks," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 7, pp. 4209–4218, 2022, doi: 10.1016/j.jksuci.2020.10.012.

[4] M. Ayaz, M. Ammad-Uddin, Z. Sharif, M. Hijji, and A. Mansour, "A hybrid data collection scheme to achieve load balancing for underwater sensor networks," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 35, no. 3, pp. 74–86, 2023, doi: 10.1016/j.jksuci.2023.02.006.

[5] A. M. Alabdali, N. Gharaei, and A. A. Mashat, "A Framework for Energy-Efficient Clustering with Utilizing Wireless Energy Balancer," *IEEE Access*, vol. 9, pp. 117823–117831, 2021, doi: 10.1109/ACCESS.2021.3107230.

[6] E. Hajian, M. R. Khayyambashi, and N. Movahhedinia, "A Mechanism for Load Balancing Routing and Virtualization Based on SDWSN for IoT Applications," *IEEE Access*, vol. 10, pp. 37457–37476, 2022, doi: 10.1109/ACCESS.2022.3164693.

[7] L. Dash *et al.*, "A Data Aggregation Approach Exploiting Spatial and Temporal Correlation among Sensor Data in Wireless Sensor Networks," *Electron.*, vol. 11, no. 7, 2022, doi: 10.3390/electronics11070989.

[8] G. P. Gupta, M. Misra, and K. Garg, "Towards scalable and load-balanced mobile agents-based data aggregation for wireless sensor networks," *Comput. Electr. Eng.*, vol. 64, pp. 262–276, 2017, doi: 10.1016/j.compeleceng.2017.10.020.

[9] A. Salim, A. Ismail, W. Osamy, and A. M. Khedr, "Compressive sensing based secure data aggregation scheme for IoT based WSN applications," *PLoS One*, vol. 16, no. 12, p. e0260634, Dec. 2021, doi: 10.1371/journal.pone.0260634.

[10] S. Math, P. Tam, and S. Kim, "Reliable federated learning systems based on intelligent resource sharing scheme for big data internet of things," *IEEE Access*, vol. 9, pp. 108091–108100, 2021, doi: 10.1109/ACCESS.2021.3101871.

[11] H. R. Farahzadi, M. Langarizadeh, M. Mirhosseini, and S. A. F. Aghda, "An improved cluster formation process in wireless sensor network to decrease energy consumption," *Wirel. Networks*, vol. 27, no. 2, pp. 1077–1087, 2021, doi: 10.1007/s11276-020-02485-y.

[12] D. Zhang, T. Zhang, J. Zhang, Y. Dong, and X. Zhang, "A kind of effective data aggregating method based on compressive sensing for wireless sensor network," *EURASIP J. Wirel. Commun. Netw.*, no. 1, p. 159, Dec. 2018, doi: 10.1186/s13638-018-1176-4.

[13] Z. Wu and G. Wan, "An enhanced ACO-based mobile sink path determination for data gathering in wireless sensor networks," *EURASIP J. Wirel. Commun. Netw.*, no. 1, p. 100, Oct. 2022, doi: 10.1186/s13638-022-02145-z.

[14] S. Bhushan, M. Kumar, P. Kumar, T. Stephan, A. Shankar, and P. Liu, "FAJIT: a fuzzy-based data aggregation technique for energy efficiency in wireless sensor network," *Complex Intell. Syst.*, vol. 7, no. 2, pp. 997–1007, 2021, doi: 10.1007/s40747-020-00258-w.

[15] R. Ahmad and A. Srivastava, "Sequential Load Balancing for Link Aggregation Enabled Heterogeneous LiFi WiFi Network," *IEEE Open J. Veh. Technol.*, vol. 3, pp. 138–148, 2022, doi: 10.1109/OJVT.2022.3165537.

[16] B. A. Begum and S. V. Nandury, "A Survey of Data Aggregation Protocols for Energy Conservation in WSN and IoT," *Wirel. Commun. Mob. Comput.*, 2022, doi: 10.1155/2022/8765335.

[17] A. Arulprakash, K. M. Baalamurugan, R. K. Dhanaraj, K. S. Kumar, P. Gupta, and S. Rehman, "Aggregation Technique Using Dynamic Cross-Propagation Clustering Algorithm in Wireless Body Sensor Networks," *Wirel. Commun. Mob. Comput.*, no. 1, Jan. 2022, doi: 10.1155/2022/6102584.

[18] J. Singh *et al.*, "Load-Balancing Strategy: Employing a Capsule Algorithm for Cutting Down Energy Consumption in Cloud Data Centers for Next Generation Wireless Systems," *Comput. Intell. Neurosci.*, no. 1, Jan. 2023, doi: 10.1155/2023/6090282.

[19] V. K. H. Prasad and S. Periyasamy, "Energy Optimization-Based Clustering Protocols in Wireless Sensor Networks and Internet of Things-Survey," *Int. J. Distrib. Sens. Networks*, 2023, doi: 10.1155/2023/1362417.

[20] X. Wang and H. Chen, "A Survey of Compressive Data Gathering in WSNs for IoTs," *Wirel. Commun. Mob. Comput.*, 2022, doi: 10.1155/2022/4490790.

[21] C. Zhang, O. Li, Y. Yang, G. Liu, and X. Tong, "Energy-efficient data gathering algorithm relying on compressive sensing in lossy WSNs," *Meas. J. Int. Meas. Confed.*, vol. 147, 2019, doi: 10.1016/j.measurement.2019.106875.

[22] M. Prabha, S. S. Darly, and B. J. Rabi, "A novel approach of hierarchical compressive sensing in wireless sensor network using block tri-diagonal matrix clustering," *Comput. Commun.*, vol. 168, pp. 54–64, 2021, doi: 10.1016/j.comcom.2020.12.017.

[23] P. Mohan *et al.*, "TPEMLB: A novel two-phase energy minimized load balancing scheme for WSN data collection with successive convex approximation using mobile sink," *Ain Shams Eng. J.*, vol. 15, no. 10, 2024, doi: 10.1016/j.asej.2024.102849.

[24] M. Mirza, V. Periyasamy, M. Ramesh, S. Mariappan, S. Rajagopal, K. Suriyan, and K. Venusamy, "Internet of thing based health monitoring system using wearable sensors networks," *International Journal of Reconfigurable and Embedded Systems (IJRES)*, vol. 13, no. 2, pp. 424–430, 2024, doi: 10.11591/ijres.v13.i2.pp424-430.

[25] S. L. Ullo and G. Sinha, "Advances in smart environment monitoring systems using IoT and sensors," *Sensors*, vol. 20, no. 11, 2020, doi: 10.3390/s20113113.

BIOGRAPHIES OF AUTHORS



Prof. Raju Chitra received the B.E. (Electronics and Communication Engineering), degree in Government College of Technology, Coimbatore in 2000. She received M.E. degree in VLSI Design from the Anna University of Technology, Coimbatore in 2010. Currently, she is working as assistant professor (SS) in Department of Electronics and Communication Engineering, School of Engineering, Avinashilingam Institute, Coimbatore and she is part-time research scholar in Department of ECE, School of Engineering, Avinashilingam Institute, Coimbatore, Tamil Nadu. She published 15 papers in international journals and presented papers in international conferences. Her research interests includes computer networks, wireless sensor networks and internet of things. She can be contacted at email: chitra_ece@avinuty.ac.in.



Dr. Rajagopal Sudarmani is currently working as professor in the Department of Electronics and Communication Engineering, School of Engineering at the Avinashilingam Institute for Home Science and Higher Education for Women, in Tamil Nadu, India. She received her B.E. Electronics and Communication Engineering from Bharathiar University, Tamil Nadu, India and her M.E. (Applied Electronics) and Ph.D. in Information and Communication Engineering from Anna University, Chennai, India. She is published 116 papers in refereed journals, conferences, and chapters in books. She also presented research papers at several national and international conferences including the International Conference IEEE TENCON 2011 at Bali, Indonesia. Her research interests includes wireless sensor networks, wireless communication, and signal processing. She can be contacted at email: sudarmani_ece@avinuty.ac.in.



Dr. Kannadhasan Suriyan is an associate professor and HoD in the Department of Electronics and Communication Engineering at Study World College of Engineering, Coimbatore, Tamil Nadu, India. He earned his Ph.D. in Smart Antennas from Anna University in 2022. He has fifteen years of teaching and research experience. He obtained his B.E. in ECE from Sethu Institute of Technology, Kariapatti, in 2009 and his M.E. in communication systems from Velammal College of Engineering and Technology, Madurai, in 2013. He obtained his M.B.A. in Human Resources Management from Tamil Nadu Open University, Chennai. He has published approximately 195 papers in reputable international journals indexed by SCI, Scopus, Web of Science, and Major Indexing and has presented or published more than 332 papers in national and international journals, book chapters, and conferences. Additionally, he has contributed a book chapter. He also serves as a board member, reviewer, speaker, session chair, and member of the advisory and technical committees of various colleges and conferences. He is also expected to participate in various workshops, seminars, conferences, faculty development programs, STTP, and online courses. His areas of interests are smart antennas, digital signal processing, wireless communication, wireless networks, embedded systems, network security, optical communication, microwave antennas, electromagnetic compatibility and interference, wireless sensor networks, digital image processing, satellite communication, cognitive radio design, and soft computing techniques. He is a member of SMIEEE, ACM, IET, ISTE, FIEI, FIETE, CSI, IAENG, SEEE, IEAE, INSC, IARDO, ISRPM, IACSIT, ICSES, SPG, SDIWC, IJSPR, GSTE, and the EAI Community. He can be contacted at email: kannadhasan.ece@gmail.com.