Weighted Boost Spearman Correlative Dual Cluster Head for Robust Transmission in Wireless Sensor Network

A. VIJAYASARATHI[1], DR. N. ELAMATHI[2]
[1] *Ph.D. Research Scholar (PT), Periyar University, Salem*
[2] *Assistant Professor. Trinity College For Women, Namakkal*

Abstract- Wireless Sensor Network (WSN) consists of tiny, low-powered sensors communicating with each other via multi-hop wireless links for efficient data transmission. Clustering of nodes is utilized in WSN to reduce number of nodes for transmitting data packets to the sink node. Several researchers have been carried out on dual cluster head selection for data transmission in WSN. In these cases, cluster member transmits data packets to Primary Cluster Head (PCH), followed by which it transmits to the Secondary Cluster Head (SCH). Finally, the SCH transmits the data packets to the base station. Despite improvement observed during data transmission, data packet delivery ratio and delay was less focused. In order to address these issues, a machine learning technique called, Weighted Boost Spearman Correlative Dual Cluster Head Selection (WB-SCDCHS) is proposedfor performing resource optimized data transmission in WSN. With the initial number of sensor nodes being setup in WSN, residual energy and memory availability of every sensor node is computed. Here, Weighted Boost Agglomerative Brown Clustering (WBABC) model is employed to form strong cluster by combining weak learner, with Hierarchical Agglomerative Brown Cluster (HABC) considered as the weak learner. The HABC groups the nodes based on the residual energy and memory availability to form clusters. The second step forms the cluster head selection performed using Spearman Rank Correlative Coefficient, where sensor node with high residual energy and bandwidth availability is selected as the cluster head. In every cluster, the sensor nodes are ranked by identifying the relationship between two variables (i.e., nodes) based on residual energy and bandwidth availability. The sensor node ranking first is considered as the primary cluster head and sensor node ranking second is considered as the secondary cluster head of that particular cluster. Finally, the secondary cluster head finds the neighboring secondary cluster head with higher bandwidth availability for transmitting data to base station (BS). This in turn helps to improve the packet delivery ratio. Experimental evaluation is conducted on factors such clustering accuracy, packet delivery ratio, delay, energy consumption and packet drop rate with respect to number of sensor nodes and packets.

Indexed Terms- Wireless Sensor Network, Primary Cluster Head, Secondary Cluster Head, Weighted Boost, Spearman Rank, Correlative Coefficient.

I. INTRODUCTION

A double cluster head (CH) approach was introduced in [1] to minimize cluster head load in large cluster. A node dormancy mechanism was introduced to balance the energy consumption of network. An optimal cluster-head function was introduced to choose CH of every cluster in each round. An optimal cluster-head function was constructed depending on residual energy and node position. However, the computational complexity gets increased with higher network data transmission and reception delays.A new energy-efficient clustering algorithm was introduced in [2] for improving energy efficiency of WSNs by minimizing and balancing the energy consumption. The lemma with dual-cluster-head mechanism was introduced to minimize the energy overhead during rotation of Cluster Heads (CHs). A non-cooperative game model was presented for balancing the energy consumption among Cluster Heads. But the designed algorithm exhibited high Spatial-temporal correlation with higher data redundancy in WSNs.

A flat and hierarchical routing scheme was introduced in [3] for maximizing the energy efficiency. The number of nodes was considered as cluster heads resulting in cluster formation. A multi-hop routing scheme was introduced to communicate with cluster head for data packets reception from all cluster members and to transmit the aggregated data along path to the sink. But flat and hierarchical routing scheme increased the delay involved in cluster formation.A dual cluster head routing protocol with super node (DCHRP4) was introduced in [4] with heterogeneity to increase the wireless sensor network lifetime. The designed protocol was used to reduce the number of cluster head selection for minimizing the energy consumption. In clustering-based sensor network cluster head transmitted the data packets to the sink node. But the packet delivery ratio was not improved by DCHRP4.

A. Motivation

Single cluster head for data transmission over WSN results in high energy consumptions and increased delay across network. This is because of maximal utilization of single cluster head in a cluster and short radio range of sensor nodes respectively. In a similar manner, data transmission via single cluster head also results in non-uniform energy consumption across the network. Hence optimal usage of cluster head leads to uniform energy consumption and improved accuracy of cluster being formed across the network. Therefore, two factors residual energy and memory availability can be introduced in a WSN in this context. Introduction of residual energy and memory availability improves the usage of sensor nodes with high residual energy and high memory availability to form similar cluster and leads to faster data transfer over the network. Next, by selection of dual cluster head based on ranking higher ranked cluster head forms the primary cluster head and data transmission is said to be performed via this node followed by which the secondary cluster head performs data transmission, therefore contributing to packet delivery ratio and reducing delay involved in data transmission in WSN over a network. Hence, the primary motivation of this work is to develop a dual cluster head selection using machine learning technique for low-delay and energy-based data transmission.

B. Contributions

The primary contributions of this paper are as follows. A novel model called, Weighted Boost Agglomerative Brown Cluster for the development of efficient data transmission in WSN over a network is first proposed. The method utilizes residual energy and memory availability of sensor nodes as criteria for node selection. A probabilistic model involving emission and transition probabilities for link addition between nodes is also developed. Subsequently, a boosting model using weight is also proposed for forming strong cluster. A dual clustering model using Spearman Correlative Coefficient with ranking separate for primary and secondary cluster is also proposed. Finally, a novel Dual Spearman Rank Correlative Cluster Head Selection algorithmis designed for low delay and high packet delivery ratio data transmission is proposed. Extensive investigations into the clustering accuracy, delay, packet delivery ratio and energy consumption are also carried out.

C. Organization of the paper

The remainder of the paper is organized as follows. Section 2 discusses related work. Section 3 describes the Weighted Boost Spearman Correlative Dual Cluster Head Selection (WB-SCDCHS) technique with the aid of block diagram and algorithm in detail. Section 4 provides the simulation setup along with the simulation parameters used for design of WB-SCDCHS. Section 5 presents the performance evaluation of the proposed method along with the comparison of state-of-the-art techniques. Finally, Section 6 concludes the paper.

II. RELATED WORKS

Sensor node clustering in WSN is employed to reduce number of nodes involved in data packet transmission to base station or sink node. Every cluster comprises of both cluster head and cluster members. The cluster member transmits the data to the cluster head. After that, the cluster head transmits the data to the base station. In this way data transmission is said to be achieved. Several research works have been done in this domain.

However, in networks involving sparse node density, the scheme is found to be insignificant with the involvement of large extent of data correlation. Hence, to improve energy efficiency in high node density WSNs, a distributed source coding (DSC) was proposed in [5] employing virtual multiple-input multiple-output (MIMO) data transmission technique in WSNs. The DSC-MIMO initially compressed redundant data via DSC and then sent the aggregated data to a virtual MIMO link, therefore contributing to energy saving. Yet another data routing technique was proposed in [6] via energy cost optimization, therefore contributing to both balanced energy consumption and efficient data transfer. An autonomous energy efficient method was presented in [7] by utilizing energy management strategy.

Existing strategies have utilized several accessing mechanisms for significant transmission due to minimum memory of nodes and hence lacked in conserving compressed data lacked in preserving compressed data authenticity. To address this issue, a compressive slender penetrative etiquette was investigated in [8] to significant data transmission in WSN with minimum interference. In [9], a route selection mechanism based on network connectivity preservation for enhancing the network lifetime was proposed where nodes were necessitated to sent different data types. The objective here was said to be attained by conserving the edge connection at the starting round in such a manner that the network connectivity would be preserved even at further rounds of communications.

One of the major issues faced in WSN is the energy consumption. Whenever the data being sensed was transmitted to the base station, the sensor node in turn consumed the energy from the battery. An algorithm called, Optimized Radio Energy Algorithm (OREA) was proposed in [10] for faster data transmission. However, the major problem dealt with optimal cluster head selection is that it makes the network service prompt. Till now, several research works have been performed in this area and several solutions were also provided. With this objective an optimal cluster head selection algorithm was designed in [11] by involving four chief criterions, namely, energy, delay, distance, and security. Optimal cluster head were selecting by hybridizing dragon fly and firefly, therefore contributing to delay and risk probability.

A data transmission mechanism based on the energy efficient data gathering technique was proposed in [12]. Here, optimal udpation of the mobile sink was performed, therefore contributing to maximum network lifetime and minimum delay. However, less concentration was made on the network lifetime. To address this issue, a three level cluster head selection procedure was proposed in [13], therefore improving the whole network routine. Despite the fact that Cluster Head (CH) is extremely congested with several nodes and dies in a swift manner due to high level of congestion by only considering the residual energy results in unequal load balance in the cluster being formed. To address this issue, an advanced clustering technique was proposed in [14] that not only considered the residual energy but also the distance between the cluster head and orientation of antenna.

A multi-tier based clustering framework was designed in [15] by employing subarea division algorithm to improve scalability and network lifetime along with data transmission in WSN. A new mechanism for electing the cluster head on the basis of the sensor physical location and its residual energy was presented in [16]. In addition different communication pattern between cluster head and base station were also analyzed depending on the transmission range and number of sensors involved in the network, therefore contributing to both network lifetime and data transmission. Yet another double cluster head mechanism employing particle swarm optimization was proposed in [17] therefore minimizing average delay.

Two-level distributing cluster head selection was made in [18] by evaluating the between member nodes and cluster head, therefore reducing the energy consumption and attaining load balancing to a greater extent. However, with the scarce utilization of resources, several nodes remained undisturbed, therefore concentrating only on the highly utilized cluster head. In [19], Krill Herd Routing optimization algorithm was presented that by way of selecting optimized path resulting in improved residual energy and transmission time significantly. Yet another optimal ant colony optimization algorithm was suggested in [20] therefore ensuring good stability specifically suited for large networks.

Henceforth, the recent research works investigated above are forcibly summarizing the requirement for a new cluster head selection technique. Thus in this work, it is focused to deliberately propose a machine learning technique which brings effective results for the enhancement of robust data transmission with high packet delivery ratio.

III. WEIGHTED BOOST SPEARMAN CORRELATIVE DUAL CLUSTER HEAD SELECTION

In this section, a resource optimized data transmission technique using machine learning called, Weighted Boost Spearman Correlative Dual Cluster Head Selection (WB-SCDCHS) is proposed. Two processes are involved in the design of WB-SCDCHS. They are cluster formation and cluster head selection. Figure 1 shows the block diagram of WB-SCDCHS method.

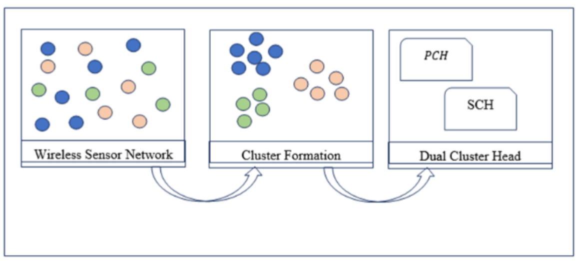

Figure 1 Block diagram of Weighted Boost Spearman Correlative Dual Cluster Head Selection

As illustrated in the above figure, prior to the launching of application (i.e., optimized data transmission in WSN), the sensor nodes are placed in strategic positions. The objective is to provide an optimized data transmission in WSN and processing performance. The formation techniques are discussed below. In the design of the network model, the total numbers of sensor nodes are denoted as 'n' with the known positioning of each node with the entire network being split into disjoint clusters. The clustering formation is based on Weighted Boost Agglomerative Brown Cluster formation.

The sensor network is denoted as a connected graph, with the set of vertices (i.e., sensor nodes '$SN = s_1, s_2, \dots, s_n$') and set of edges (i.e., transmission links 'l_i, l_j') with the predefined number of base stations (i.e., 'BS') . The objective here remains in identifying strong cluster by combining weak learner and maximizing one-hop connectivity to reduce the consumed energy in the network and improving the clustering accuracy. Furthermore, Dual Spearman Rank Correlative Cluster Head Selection is utilized for selecting cluster head based on two factors, residual energy and bandwidth. The set of experiments involves different quantities of sensor nodes and packets for computing the clustering accuracy, delay, packet delivery ratio and energy consumption.

3.1 Weighted Boost Agglomerative Brown Cluster formation

In this section, cluster formation using Weighted Boost Agglomerative Brown Cluster is presented. Initially, number of sensor nodes are obtained to form a wireless sensor network. Next, residual energy and memory availability of every sensor node is evaluated. Weighted Emphasis Boost Clustering is carried out to form strong cluster by combining weak learners for resource optimized data transmission. In the designed technique, Hierarchical Agglomerative Brown Cluster (HABC) is considered as the weak learner.

The HABC groups the sensor nodes based on the residual energy and memory availability to form clusters. The results of weak learners are ensembled to form strong cluster with significant clustering accuracy. Figure 2 given below shows the block diagram of Weighted Boost Agglomerative Brown Cluster formation model.

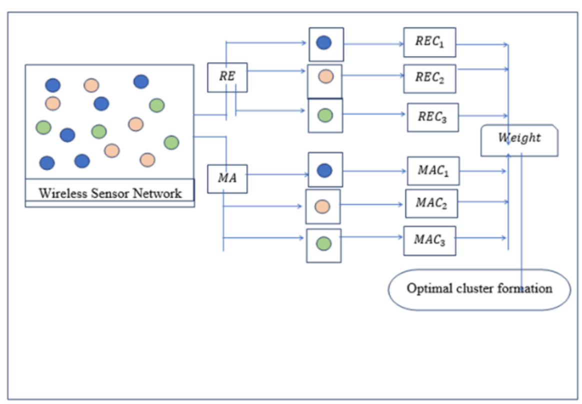

Figure 2 Block diagram of Weighted Boost Agglomerative Brown Cluster formation model

As shown in the above figure, initially, the residual energy of all the sensor node involved in the network is measured, with residual energy referring to the remaining energy. To start with, all the sensor node in the network possess similar energy. Due to continuous monitoring and sensing, the energy level gets reduced and the residual energy is mathematically formulated as given below.

$$RE(SN) = \sum_{i=1}^{n} TE(SN_i) - CE(SN_i) \qquad (1)$$

$$MA(SN) = \sum_{i=1}^{n} TM(SN_i) - CM(SN_i) \qquad (2)$$

From the above equations (1) and (2), the residual energy 'RE' and memory availability 'MA' of each sensor node 'SN' is evaluated based on the total energy 'TE', consumed energy 'CE', total memory 'TM' and consumed memory 'CM' respectively. With the above obtained residual energy 'RE' and memory availability 'MA', by employing Hierarchical Agglomerative Brown Cluster, weak learners are first obtained to form clusters. Let us consider 'v' being the set of all sensor nodes '$SN = SN_1, SN_2, \ldots, SN_n$' with residual energy '$RE$' and memory availability 'MA' in the entire network, '$C: v \rightarrow \{1,2,..,k\}$' denotes the partition of the different sensing nodes (i.e. denoted by three different colors, blue, green orange) into 'k' classes. Then, the Hierarchical Agglomerative Brown Cluster forming weak learning is mathematically expressed as given below.

$$p(SN_1, SN_2, \ldots, SN_n) = \prod_{i=1}^{n} e\left(SN_i | C(SN_i)\right) q\left(C(SN_i) | C(SN_{i-1})\right) \qquad (3)$$

From the above equation (3), 'C' corresponds to the partition of different sensing nodes with '$C(SN_i)$' denoting the starting state, '$e\left(SN_i | C(SN_i)\right)$' representing the emission probabilities and '$q\left(C(SN_i) | C(SN_{i-1})\right)$' denoting the transition probability respectively. Finally, cluster formation is performed by Weighted Emphasis Boost Clustering where the weight of each edge corresponds to the Euclidean distance between two sensor nodes. Let us select an edge 'SN_i, SN_j' with minimum weight such that '$u \in SN_i$' and '$v \notin SN_i$', then cluster formation is made by employing weight as given below for each weak learner obtained from (3).

$$Dis\left(u, v_j\right) = \sqrt{\sum_{i=1}^{n} \left(v_j - u_i\right)^2} \qquad (4)$$

The pseudo code representation of Weighted Boost Agglomerative Brown Cluster formation is given below.

Input: Sensor Nodes '$SN = SN_1, SN_2, \ldots, SN_n$'
Output: Optimal cluster '$CL = cl_1, cl_2, \ldots, cl_3$'
Step 1: Initialize total energy 'TE', consumed energy 'CE' Step 2: Begin Step 3: For each sensor nodes 'SN' and edge 'u', 'v' Step 4: Measure residual energy using equation () Step 5: Measure memory availability using equation () Step 6: Obtain weak class learner using equation (3) Step 7: Evaluate weight according to Euclidean distance using (4) Step 8: Choose an edge 'u', 'v' with minimal weight Step 9: Return (optimal cluster) Step 10: End for Step 11: End

Algorithm 1 Weighted Boost Agglomerative Brown Cluster formation

As given in the above Weighted Boost Agglomerative Brown Cluster formation algorithm, the objective remains in forming cluster with minimum energy consumption and maximum clustering accuracy. These two objectives are said to be achieved by evaluating the residual energy and memory availability of each sensor node. Based on it Agglomerative Brown Clusters weak learners are first formed, therefore minimizing the energy consumption involved in clustering and then combined using Weighted Boost to form strong clustering, therefore improving the clustering accuracy.

3.2 Dual Spearman Rank Correlative Cluster Head Selection

Upon successful cluster formation, every cluster is controlled by the cluster head. In our work dual cluster head selection is carried out using Spearman Rank Correlative Coefficient. Here, the sensor node with higher residual energy and bandwidth availability is selected as the cluster head. In every cluster, the sensor nodes are ranked according to the relationship between two variables (i.e., nodes) based on residual energy and bandwidth availability. Figure 3 shows the block diagram of Dual Spearman Rank Correlative Cluster Head Selection model.

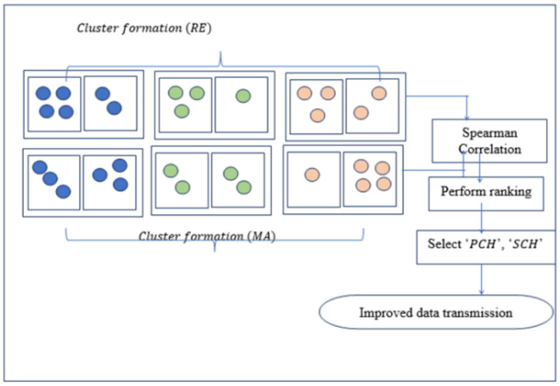

Figure 3 Block diagram of Dual Spearman Rank Correlative Cluster Head Selection model

As shown in the above figure, for a sample size of 'n' sensor nodes, the optimal clusters 'cl_i', 'cl_i' are converted into ranks 'rcl_i', 'rcl_j' with which the spearman rank is mathematically expressed as given below.

$$r_s = \rho rcl_i rcl_j = \frac{cov(rcl_i rcl_j)}{\sigma rcl_i \sigma rcl_j} \quad (5)$$

From the above equation (5), 'ρ' corresponds to the correlation coefficient, '$rcl_i rcl_j$' represents the covariance of rank clusters 'cl_i', 'cl_j' and '$\sigma rcl_i \sigma rcl_j$' represents the standard deviation rank clusters 'cl_i', 'cl_j' respectively. Only if all 'n' sensor nodes and 'n' clusters are distinct, the above equation (5) is mathematically formulated as given below.

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_i}{n(n^2-1)} \quad (6)$$

$$d_i = r(cl_i) - r(cl_j) \quad (7)$$

From the above equation (6) and (7), 'r_s' is measured for distinct rank and 'd_i' corresponds to the difference between the two ranks of each cluster. The sensor node ranking first is considered as the primary cluster head and the sensor node ranking second is considered as the secondary cluster head of that particular cluster. The sensor node ranking first is considered as the primary cluster head and the sensor node ranking second is considered as the secondary cluster head of that particular cluster.

All the nodes in the cluster transmit the data to the primary cluster head through neighboring higher bandwidth availability sensor nodes. After that, the primary cluster head transmits the data to the secondary cluster head in same cluster. Finally, the secondary cluster head finds the neighboring secondary cluster head with higher bandwidth availability for transmitting data to the base station (BS). This in turn helps to improve the packet delivery ratio. The pseudo code representation of Dual Spearman Rank Correlative Cluster Head Selection is given below.

Input: Optimal clusters '$CL = cl_1, cl_2, \dots, cl_3$'
Output: Robust Cluster Head Selection
Step 1: Initialize 'n' sensor nodes, correlation coefficient '$\rho = between + 1\ and - 1$' Step 2: Begin Step 3: For each optimal cluster '$CL = cl_1, cl_2, \dots, cl_3$' Step 4: Measure spearman rank using equation (5) Step 5: Measure spearman correlation coefficient for distinct rank using equation (6) Step 6: Return cluster head Step 7: End for Step 8: End

Algorithm 2 Dual Spearman Rank Correlative Cluster Head Selection

As given in the above Dual Spearman Rank Correlative Cluster Head Selection algorithm, the objective remains in selecting the cluster head with minimum delay and maximum packet delivery ratio, therefore ensuring efficient data transmission in WSN. To achieve this objective, a dual cluster head selection model is designed based on two factors, residual energy and bandwidth (i.e. data packet transmission over a specific cluster head in a cluster for a given amount of time).

The sensor with maximum residual energy and bandwidth is initially computed, therefore minimizing the delay involved in selecting the cluster head. Then, based on Spearman Ranking, higher rank nodes are selected as the primary cluster head. Followed by which the next higher rank node is selected as the secondary cluster head. In this manner, dual cluster head selection is made with higher bandwidth availability for transmitting data packet to the base station, therefore improving packet delivery ratio.

IV. EXPERIMENTAL SETUP

In this section, our technique and two other techniques are evaluated by simulating the technique on the NS2 simulator. To start with, 500 sensor nodes is independently and uniformly deployed in an optimal pattern with a communication radius of 30m. Then, we simulate our data transmission technique and two other techniques to compare the differences between them with different network parameters.

We adopt the following evaluation indicators for the performance metrics: clustering accuracy, packet delivery ratio, energy consumption and delay. When building the network model, all the wireless sensor nodes are randomly distributed with a network area of 100 * 100 m^2 and the parameters in the simulation is provided in table 1.

Table 1 Simulation parameters

S. No	Parameter	Values
1	Network area	100 * 100 m^2
2	Number of nodes	500
3	Initial energy	0.5 J
4	Packet	15/iteration
5	Bandwidth	10 bit/s
6	Data packet size	2000 bits
7	Communication range	30m

4.1 Implementation scenario

In this section, the implementation scenario for efficient data transmission in WSN via cluster formation and cluster head selection is elaborated in brief. To start with cluster formation is performed by employing Weighted Boost Agglomerative Brown Cluster. The implementation scenario in this section is performed via 15 different numbers of sensor nodes, each sensing different regions differentiated by means of three different colors '$Blue\ [B] - 6, Green\ [G] - 4\ and\ Orange\ [O] - 5$' respectively. Table 2 given below provides the residual energy measurement for 15 different sensors by applying equation (1).

Table 2 Residual energy measurement

Sensor nodes	Initial energy	Consumed energy	Residual energy
$sn_1[B]$	$0.5J$	$0.15J$	$0.35J$
$sn_2[G]$	$0.5J$	$0.25J$	$0.25J$
$sn_3[B]$	$0.5J$	$0.15J$	$0.35J$
$sn_4[O]$	$0.5J$	$0.30J$	$0.20J$
$sn_5[O]$	$0.5J$	$0.10J$	$0.40J$
$sn_6[O]$	$0.5J$	$0.30J$	$0.20J$
$sn_7[B]$	$0.5J$	$0.15J$	$0.35J$
$sn_8[O]$	$0.5J$	$0.30J$	$0.20J$
$sn_9[B]$	$0.5J$	$0.35J$	$0.15J$
$sn_{10}[G]$	$0.5J$	$0.25J$	$0.25J$
$sn_{11}[G]$	$0.5J$	$0.20J$	$0.30J$
$sn_{12}[B]$	$0.5J$	$0.15J$	$0.35J$
$sn_{13}[O]$	$0.5J$	$0.10J$	$0.40J$
$sn_{14}[G]$	$0.5J$	$0.25J$	$0.25J$
$sn_{15}[B]$	$0.5J$	$0.35J$	$0.15J$

Table 3 given below provides the memory availability measurement for 15 different sensors by applying equation (2).

Table 3 Memory availability measurement

Sensor nodes	Initial memory	Consumed memory	Memory available
$\boldsymbol{sn_1[B]}$	$2000MB$	$700MB$	$1300MB$
$\boldsymbol{sn_2[G]}$	$2000MB$	$1000MB$	$1000MB$
$\boldsymbol{sn_3[B]}$	$2000MB$	$800MB$	$1200MB$
$\boldsymbol{sn_4[O]}$	$2000MB$	$700MB$	$1300MB$
$\boldsymbol{sn_5[O]}$	$2000MB$	$700MB$	$1300MB$
$\boldsymbol{sn_6[O]}$	$2000MB$	$700MB$	$1300MB$
$\boldsymbol{sn_7[B]}$	$2000MB$	$700MB$	$1300MB$
$\boldsymbol{sn_8[O]}$	$2000MB$	$700MB$	$1300MB$
$\boldsymbol{sn_9[B]}$	$2000MB$	$800MB$	$1200MB$
$\boldsymbol{sn_{10}[G]}$	$2000MB$	$900MB$	$1100MB$
$\boldsymbol{sn_{11}[G]}$	$2000MB$	$900MB$	$1100MB$
$\boldsymbol{sn_{12}[B]}$	$2000MB$	$700MB$	$1300MB$
$\boldsymbol{sn_{13}[O]}$	$2000MB$	$800MB$	$1200MB$
$\boldsymbol{sn_{14}[G]}$	$2000MB$	$1000MB$	$1000MB$
$\boldsymbol{sn_{15}[B]}$	$2000MB$	$800MB$	$1200MB$

Then by applying Hierarchical Agglomerative Brown Cluster are given in equation (3) is obtained as given below.

$$Residual\ Energy - eB(0.35) = 4; eB(0.15) = 2; eG(0.25) = 3; eG(0.30) = 1; eO(0.35) = 3; eO(0.15) = 2$$

$$Memory\ Availability - qB(1300) = 3; qB(1200) = 3; qG(1000) = 2; qG(1100) = 2; qO(1200) = 1; qO(1300) = 4$$

The results of the above clustering evaluation are illustrated in figure 3. Then, the results of spearman rank using equations (5), (6) and (7) are given below in table 4.

TABLE 4 DUAL CLUSTER HEAD SELECTION

Sensor nodes	Residual energy	Rank 1	Adjust rank 1	Memory available	Rank 2	Adjust rank 2	d	d^2
1	$0.35J$	3	4.5	1300.	1	1	3	9
2	$0.25J$	8	9	1000.	14	19.5	−10	110.
3	$0.35J$	3	4.5	1200.	8	9.5	−5	25
4	$0.20J$	11	12	1300.	1	1	11	121
5	$0.40J$	1	1.5	1300.	1	1	.5	25
6	$0.20J$	11	12	1300.	1	1	11	121
7	$0.35J$	3	4.5	1300.	1	1	3.5	12.2
8	$0.20J$	11	12	1300.	1	1	11	121
9	$0.15J$	14	19.5	1200.	9.5	9.5	10	100
10	$0.25J$	8	9	1100.	12.5	12.5	−3.5	12.2
11	$0.30J$	7	7	1100.	12.5	12.5	−5.5	30.2
12	$0.35J$	3	4.5	1300.	1	1	3.5	12.2
13	$0.40J$	1	1.5	1200.	9.5	9.5	−8	64
14	$0.25J$	8	9	1000.	19.5	19.5	−10	110.
15	$0.15J$	14	19.5	1200.	9.5	9.5	10	100

Finally, the results of the cluster head selection based on the spearman ranking is given below in figure 4.

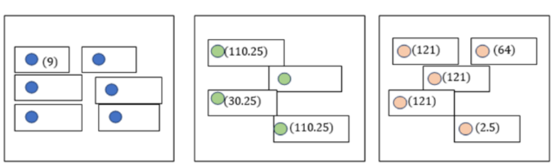

Figure 4 Cluster head selection

As formed from the three different clusters, the node with higher values is selected as the primary cluster head and the second node as the secondary cluster head, followed by which the data transmission is performed to the base station.

4.2 Performance evaluation and discussion

In this section the performance evaluation of three different parameters with detailed comparison between the proposed WB-SCDCHS and existing Double Cluster Head [1] and Energy-efficient Clustering [2] is given below. For fair comparison similar number of sensor nodes and parameters settings is utilized for three different techniques.

4.2.1 Case 1: Clustering Accuracy

The first metric of significance for robust transmission in wireless sensor network is the clustering accuracy. Higher the accurate clusters being formed, more efficient the cluster head being selected and therefore higher is the data transmission rate. The mathematical formulation of clustering accuracy is measured as given below.

$$CA = \frac{C_{af}}{TC} * 100 \qquad (8)$$

From the above equation (8), the clustering accuracy 'CA' is measured on the basis of the cluster accurately being formed 'C_{af}' to the total clusters 'TC' in the network for the corresponding number of sensor nodes 'SN'. It is measured in terms of percentage (%). Table 5 given below shows the performance measure of clustering accuracy for three different techniques.

Table 5 Performance measure of clustering accuracy

Sensor nodes	Clustering accuracy (%)		
	WB-SCDCHS	Double Cluster Head	Energy-efficient Clustering
50	80	70	60
100	79.35	68.55	59.35
150	78.45	68.35	59
200	77.35	68	58.75
250	76.55	67.45	58.35
300	75	67.25	58
350	74.25	67	57.35
400	73	66.55	57.25
450	71.45	66.35	57
500	70	66	56.55

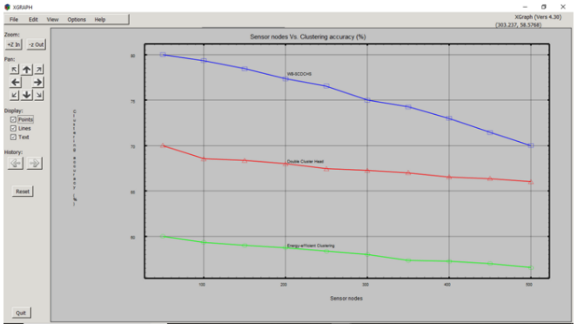

Figure 5 NS2 graphical representation of clustering accuracy

Figure 5 given above shows the clustering accuracy for different numbers of sensor nodes in the range of 50 to 500 for a network area of 100 * 100 m^2. From the figure it is inferred that the clustering accuracy is inversely proportional to the number of sensor nodes considered for simulations. In other words, increasing the number of sensor nodes causes a decrease in the clustering accuracy. This is due to the reason that with the increase in the sensor nodes deployed in WSN, during the formation of cluster and cluster head variations may said to occur and with this, the accuracy is also said to be reduced. However, with '50' sensor nodes considered for simulation and correct number of clusters being '10', the cluster accurately formed using WB-SCDCHS was found to be '8', contributing to clustering accuracy as '80%'. In a similar manner using [2] and [2], the cluster accurately formed was found to be '7' and '6' respectively, therefore resulting in clustering accuracy to be '70%' and '60%' respectively. From the results we can say that the clustering accuracy using WB-SCDCHS is said to be comparatively better than [1] and [2]. The reason behind the improvement is due to the application of Weighted Boost Agglomerative Brown Cluster formation algorithm. By applying this algorithm the Clusters are said to be formed based on the residual energy and memory availability of each sensor node. So, sensor node possessing high residual energy and memory are used as the basis for forming strong cluster, therefore improving the clustering accuracy using WB-SCDCHS by 12% compared to [1] and 30% compared to [2].

4.2.2 Case 2: Energy Consumption

The second metric to be of significance for effective data transmission in WSN is the energy consumption. The energy consumption refers to the energy being consumed during clustering and is mathematically formulated as given below.

$$EC = \sum_{i=1}^{n} sn_i * EC[p(SN_1, SN_2, \dots, SN_n)] \quad (9)$$

From the above equation (9), the energy consumption 'EC' is measured based on the number of sensor nodes considered for simulation 'sn_i' and the energy consumed in forming the cluster according to the partitioning of different sensing nodes '$EC[p(SN_1, SN_2, \dots, SN_n)]$'. It is measured in terms of joules 'J'. Table 6 given below shows the performance measure of energy consumption for three different techniques.

Table 6 Performance measure of energy consumption

Sensor nodes	Energy Consumption (J)		
	WB-SCDCHS	Double Cluster Head	Energy-efficient Clustering
50	12.5	16	19
100	13.2	17.5	21.5
150	14.5	18.3	24.5
200	16.5	20.5	25
250	19	23.5	27.4
300	21.5	26.8	33.5
350	23	30	38.4
400	24.5	33.5	42.3
450	25	35	45
500	27.2	40	50

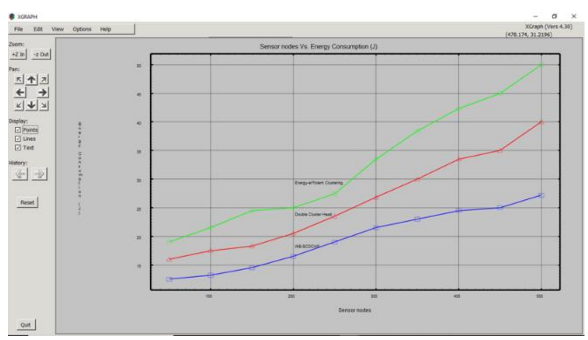

Figure 6 NS2 Graphical representation of energy consumption

Figure 6 given above shows the graphical representation of energy consumption involved in clustering for three different techniques with respect to 500 different numbers of sensor nodes for an initial energy setting of '$0.5J$'. From the figure it is inferred that increasing the number of sensor nodes causes an increase in the number of nodes involved in clustering and therefore results in the increasing energy consumption. However, with '50' numbers of sensor nodes considered for simulation and the energy consumed for partitioning the sensor nodes being '$0.25J$' using WB-SCDCHS, '$0.32J$' using [1] and '$0.38J$' using [2], the overall energy consumption was observed to be '$12.5J$', '$16J$' and '$19J$' respectively. From this the energy consumption is found to be minimum using WB-SCDCHS than [1] and [2]. The reason behind the improvement is due to the incorporation of emission probabilities and transition probabilities being measured separately for residual energy and memory availability. Based on these measures, clusters were formed and therefore, the energy consumption using WB-SCDCHS was found to be lesser than 24% compared to [1] and 39% compared to [2].

4.2.3 Case 3: Packet Delivery Ratio

Finally, to measure the robust transmission, the packet delivery ratio has to be measured. The packet delivery ratio refers to the percentage ratio of data packets transferred from the source to the data packets received at the base station and is mathematically formulated as given below.

$$PDR = \frac{DP_{received}}{DP_{sent}} * 100 \qquad (10)$$

From the above equation (10), the packet delivery ratio 'PDR' is measured based on the data packets received '$DP_{received}$' and the data packet sent 'DP_{sent}'. It is measured in terms of percentage (%). Table 7 given below shows the performance measure of energy consumption for three different techniques.

Table 7 Performance measure of packet delivery ratio

Data Packets	Packet Delivery Ratio (%)		
	WB-SCDCHS	Double Cluster Head	Energy-efficient Clustering
15	80	73.33	66.66
30	83.25	75.45	68.95
45	85	78	70.25
60	82.15	74.35	68.35
75	82.45	74	68
90	81.25	72.15	66.35
105	80	70.45	65
120	78.35	69.55	66.25
135	81.45	72.35	67
150	83	74	68

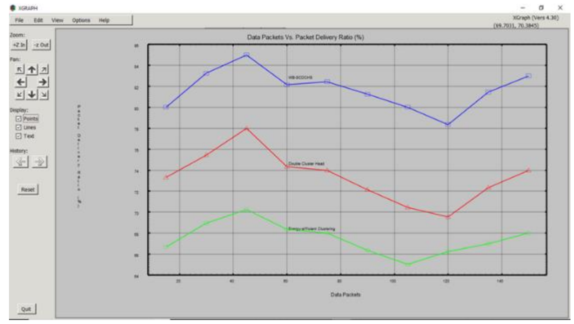

Figure 7 NS2 Graphical representation of packet delivery ratio

Figure 7 given above shows the packet delivery ratio of three different techniques with respect to 150 different packets techniques with 15 packets per iteration. Ten different simulation runs are performed to measure the packet delivery ratio. In addition, from the figure it is inferred that the packet delivery ratio neither increases not decreases with the increase in the number of data packets. This is because of the frequent topology changes due to the dynamic nature of the network and also the frequent movement of the nodes from one cluster to another that causes a variation in the packet delivery ratio. With '15' packets considered for simulation and '12' packets received correctly at the base station, the packet delivery ratio using WB-SCDCHS was observed to be '80%', with '11' and '10' packets received at the base station using [1] and [2], the packet delivery ratio was observed to be '73.33%' and '66.66%' respectively. From the results it is inferred that the packet delivery ratio using WB-SCDCHS is comparatively better than [1] and [2]. The reason behind the improvement is the application of Dual Spearman Rank Correlative Cluster Head Selection algorithm. By applying this algorithm, using Spearman Rank Correlative Coefficient, dual clusters are selected based on two factors, residual energy and bandwidth. With this only higher-ranking sensor node is selected as the primary cluster head for data transmission and therefore, the packet delivery ratio using WB-SCDCHS is found to be improved by 11% compared to [1] and 21% compared to [2].

CONCLUSION

Dual cluster head selection is a proven technique for robust data transmission in WSNs. The state-of-the-art solutions in this research area mostly focused on the cluster head selection based on the optimization distance between the nodes. The Weighted Boost Spearman Correlative Dual Cluster Head Selection (WB-SCDCHS) technique proposed in this paper aims to increase the clustering accuracy and packet delivery ratio, therefore ensuring robust data transmission between sensor nodes and base station via cluster head along with minimum energy consumption. WB-SCDCHS as an agglomerative brown clustering algorithm forms the clusters in an iterative manner based on two factors, residual energy and memory availability of each sensor node and then applying weighted boosting to form strong clusters. In addition, the idea of choosing the cluster head was also applied in the proposed WB-SCDCHS technique based on dual cluster head via machine learning. This idea reduced the delay and improved packet delivery ratio. The extensive simulations are done with different sensor nodes to compare the performance of WB-SCDCHS technique with the state-of-the-art clustering solutions including Double Cluster Head and Energy-efficient Clustering. The collected results show that the WB-SCDCHS outperforms in average the other algorithms in terms of clustering accuracy criterion with 12% and 30% compared to Double Cluster Head and Energy-efficient Clustering. With regard to energy consumption, the proposed technique saved 24% and 39% compared to Double Cluster Head and Energy-efficient Clustering. The results showed that the proposed WB-SCDCHS technique succeeded in minimizing energy consumption via effective dual cluster head selection and consequently, increased the data transmission of WSN.

REFERENCES

[1] Zhidong Zhao, Duoshui Shi, Guohua Hui, and Xiaohong Zhang, "An Energy-Optimization Clustering Routing Protocol Based on Dynamic Hierarchical Clustering in 3D WSNs", IEEE Access, Volume 7, June 2019, Pages 80159 – 80173

[2] Deyu Lin and Quan Wang, "An Energy-Efficient Clustering Algorithm Combined Game Theory and Dual- Cluster-Head Mechanism for WSNs", IEEE Access, April 2019, Volume 7, Pages 49894 – 49905

[3] Mohit Sajwan, Devashish Gosain and Ajay K. Sharma, "Hybrid energy-efficient multi-path routing for wireless sensor networks", Computers and Electrical Engineering, Elsevier, Volume 67, 2018, Pages 96–113

[4] YashaIstwal and Shashi Kant Verma, "Dual Cluster Head Routing Protocol with Super Node in WSN", Wireless Personal Communications, Springer, Volume 104, 2019, Pages 561–575

[5] Manish Kumar Singh, Syed Intekhab Amin, "Energy-efficient data transmission technique for wireless sensor networks based on DSC and virtual MIMO", Wiley, Jan 2020

[6] Om Jee Pandey, Rajesh M. Hegde, "Low-Latency and Energy-Balanced Data Transmission over Cognitive Small World WSN", IEEE Transactions on Vehicular Technology (Volume: 67, Issue: 8, Aug. 2018)

[7] C.S. Abella, S. Bonina, A. Cucuccio, S. D'Angelo, G. Giustolisi, A.D. Grasso, A. Imbruglia, G.S. Mauro, G.A.M. Nastasi, G. Palumbo, S. Pennisi, G. Sorbello, A. Scuderi," Autonomous Energy-Efficient Wireless Sensor Network Platform for Home/Office Automation", IEEE Sensors Journal (Volume: 19, Issue: 9, May1, 1 2019)

[8] Vandana Reddy, P. Gayathri, "Energy efcient data transmission in WSN thru compressive slender penetrative etiquette", Journal of Ambient Intelligence and Humanized Computing, Springer, Mar 2020

[9] Aarti Jain, Manju Khari, Elena Verdu, Shigeru Omatsu, Ruben Gonzalez Crespo, "A route selection approach for variable data transmission in wireless sensor networks", Cluster Computing, Springer, Apr 2020

[10] S. Senthil Kumar, J. T. Thirukrishna, "An Efcient QoS Based Data Packet Transmission in Wireless Sensor Networks Using OREA", Wireless Personal Communications, Springer, Apr 2020

[11] Turki Ali Alghamdi, “Energy efficient protocol in wireless sensor network: optimized cluster head selection model”, Telecommunication Systems, Springer, Mar 2020
[12] R. Ashween, B. Ramakrishnan, M. Milton Joe, “Energy Efcient Data Gathering Technique Based on Optimal Mobile Sink Node Selection for Improved Network Life Time in Wireless Sensor Network (WSN)”, Wireless Personal Communications, Springer, Apr 2020
[13] Rahul Priyadarshi, Piyush Rawat, Vijay Nath, Bibhudendra Acharya, N. Shylashree, “Three level heterogeneous clustering protocol for wireless sensor network”, Microsystem Technologies, Springer, May 2020
[14] Malek Alrashidi, Nejah Nasri, Salim Khediri, Abdennaceur Kachouri, “Energy-Efciency Clustering and Data Collection for Wireless Sensor Networks in Industry 4.0”, Journal of Ambient Intelligence and Humanized Computing, Springer, May 2020
[15] Anurag Shukla, Sarsij Tripathi, “A multi-tier-based clustering framework for scalable and energy efficient WSN-assisted IoT network”, Wireless Network, Springer, Feb 2020
[16] Tariq Taleb, Mejdi Kaddour, “Hierarchical Agglomerative Clustering Schemes for Energy-efficient in Wireless Sensor Networks”, Transport and Telecommunication, Jul 2017
[17] T. Preethiya, A. Muthukumar, S. Durairaj, “Double Cluster Head Heterogeneous Clustering for Optimization in Hybrid Wireless Sensor Network”, Wireless Personal Communications, Springer, Sep 2019
[18] S. M. Amini, A. Karimi, “Two-level distributed clustering routing algorithm based on unequal clusters for large-scale Internet of Things networks”, The Journal of Supercomputing, Springer, Nov 2019
[19] P. Visu, T. Suriya Praba, N. Sivakumar, R. Srinivasan, T. Sethukarasi, “Bio-inspired dual cluster heads optimized routing algorithm for wireless sensor networks”, Journal of Ambient Intelligence and Humanized Computing, Springer, Jan 2020
[20] Jinyu Ma, Shubin Wang, Chen Meng, Yanhong Ge, Jingtao Du, “Hybrid energy-efficient APTEEN protocol based on ant colony algorithm in wireless sensor network”, EURASIP Journal on Wireless Communications and Networking, Jul 2018