

An Energy-Efficient Fuzzy Based Data Fusion and Tree Based Clustering Algorithm for Wireless Sensor Networks

Veeramuthu Venkatesh¹(✉), Pethuru Raj², and P. Balakrishnan³

¹ School of Computing, SASTRA University, Thanjavur 613401, India
veeramuthuvenkatesh@cse.sastra.edu

² Reliance Jio Cloud Services (JCS), Bangalore 560025, India

³ SCOPE, Department of Analytics, VIT University Vellore Campus,
Vellore, India

Abstract. The realization of any wireless sensor network (WSNs) clearly determined on how the key quality of service (QoS) attributes/non-functional requirements (NFRs) gets accomplished. The well-known issues to be taken into account while designing WSNs for setting up smarter environments are information precision (timeliness and accuracy), data accretion, network latency, and energy efficiency. A wisely employed clustering algorithm for interconnecting sensor nodes can significantly enhance the energy efficiency of WSNs. However, the aspect of clustering involves additional overheads due to cluster head selection and cluster construction. This research work proposes a workaround that utilizes Type-II fuzzy for fusing the data and tree driven clustering algorithm that employs Type-2 fuzzy logic to improve the QoS parameters as well as preserving the power/energy of sensor networks. The primary objectives of the proposed cluster algorithm are two folded. Firstly, it constructs the clusters and chooses the cluster head (CH) by considering the remaining energy in the nodes and its distance from the base station (BS). Secondly, it performs the data fusion which contains meaningful information that has been sensed and captured. An extensive experimental analysis has been done on the proposed FBDF-TBC method by comparing it against its counterparts. The simulation results conclude that the proposed fuzzy-based technique for data fusion and tree-based clustering routing algorithm (FBDF-TBC) outperforms other clustering algorithms and improves the overall network lifetime of WSN from a minimum of 16% to maximum of 76%.

Keywords: Mobile sensor networks · Routing protocol · Network lifetime · Fuzzy based clustering · Delay · Data fusion

1 Introduction

The wireless sensor networks (WSNs) are gaining unprecedented popularity due to the distinct advantages that they bring to the table. That is, sensors emerge as a low-cost alternative for a wider variety of application areas across multiple industry verticals [1]. The application domains of WSNs are growing consistently with researchers unearthing

novel use cases. The well-known application areas are pattern/activity/gesture recognition, environmental monitoring, event-driven applications, self, surroundings and situation awareness, safety, surveillance and security applications, the formation of smart environments such as smarter homes, hotels, and hospitals, and so on. Primarily there are two types of WSNs: homogeneous and heterogeneous. An apostolic structure of several sensors are contemplated in these two varieties of WSNs. The lifetime and the dependability of any sensor networks can be enhanced by the heterogeneity aspect. Heterogeneous sensor networks (HSNs) are highly beneficial because they are very right and relevant to many real-world and life scenarios [2]. However, the task of routing has been a challenging concern in the design of WSNs. However, the earlier research works have come out with a few pioneering routing protocols to reduce the energy consumption by nodes in order to enhance WSN routing [3–6]. The concept of clustering is pronounced as an important factor to substantially increase the lifetime of WSNs as clusters are typically able to significantly decrease energy consumption [7]. Clusters come handy in setting and sustaining scalable sensor networks in order to tackle more data loads. The availability of sensor networks is guaranteed through clustering. Thus the concept of clustering of various participating and contributing nodes is being termed as the most important domain for intense study and research. The proven master and slave concept is doing well in forming sensor networks. There have to be one or more master nodes in order to keep the slave nodes well. If there is any fallout or problem with worker nodes, the master node has to do the necessary corrective actions immediately in order to finish the work started. The master node is typically touted as the cluster head (CH). The cluster head is nominated and designated by all the sensors in the cluster. That can be also decided by network designer. The traditional clustering protocols in WSNs assume that every nodes in the sensor networks remain stuffed with the same amount of energy. This assumption comes in the way of leveraging the extreme benefits of heterogeneous nodes and networks. In order to use the heterogeneity, clustering procedures are mainly classified based on two main benchmarks according to its meta-stability and energy-efficiency factors. The selection of cluster head for deriving energy-efficient networks generally depends on the early energy, residual energy and intermediate energy of the network and the energy depletion rate or the mixture of these parameters. The chosen protocols for clustered HSNs prolong the time intermission before the death of first node. This is called the meta-stability period.

1.1 The Key Contributions of the Proposed Framework Are Foregrounded as Follows

Tree-based clustering algorithm (TBC) for effective cluster formation. The limited battery power of each node is a major factor capable of adversely impacting the lifetime of the entire network. Tree-based [3–7] thereby it is a good ploy for extending the network lifetime. While performing tree-based clustering, there is a need to smartly construct clusters to decrease the communication distance among the sensor nodes.

Type-2 Fuzzy Logic Based Sensor Data Fusion (FBDF) for the Removal of Redundant Information Thereby Improving the Energy. Sensor Data Fusion Removes Any Incorrect and Duplicated Values from Sensors to Increase the Intended QoS. There Are Many Fusion Techniques [8, 9] Available to Perform Data Fusion but Type-2 Fuzzy Logic Provides a Greater Accuracy and Reduces Energy Utilization in Sensing the Environment

Distributed Source Coding for compression. In addition, the compression of data before transmitting greatly reduces the energy consumption by decreasing the number of bits to be transferred. In the case of sensor nodes, the common compression technique such as Huffman is not suitable as it requires an enormous amount of memory and robust processing element capability. Distributed Source Coding (DSC) [10] method by Slepian-Wolf theorem provides a precise restoration of data for twofold associated sources using side data sources. Therefore DSC is decided as the most appropriate compression technique for WSN so as to save energy.

The following part of paper is described as: In Sect. 2, a study on existing methods are described. The framework and the problem summary are briefed in Sect. 3 followed by Sect. 4 that elucidates the proposed framework of fuzzy based data-fusion technique. Section 5 elucidates the brief analysis of the performance of the framework. At last, the conclusion and future enhancements is summarized in Sect. 6.

1.2 Related Work

Firstly, this research work supplies the detailed literature survey on any energy aware in wireless sensor networks and groups them based on their objectives (Cluster formation, Cluster head selection, fault-tolerance, packet delay, Sub-clustering, and compression). Secondly, the literature surveyed clearly identifies the gaps, articulates the objectives of the proposed work and carefully formulate the solution methodologies for these objectives. In [11] proposed a fuzzy-based unequal clustering procedure in WSNs to generate clusters with different sizes and this arrangement addresses the persistent hotspot problem. As a result, the paper claims that their method decreases the intra-cluster functions of the cluster-heads which are close to the base station or have low residual energy. In [12] proposed a new methodology to construct a data gathering with energy efficiency as main motto in wireless mobile sensor networks. Subsequently, the paper concludes that the suggested approach minimizes the delay per round and guarantees an improved throughput, and eliminates minimum coverage cost of any underlying network. In [13] have come out with an architecture for any wireless micro sensor networks which are application-specific protocol that includes low-energy adaptive clustering hierarchy (LEACH) to combine the benefits of energy efficiency and lifetime and media access to attain better network lifetime, response time, and application- comprehended value. In [14] proposed a three fuzzy descriptors method based cluster-head election for wireless sensor networks which is more suitable for medium sized clusters. However, the authors highlight that the articulated model introduces a substantial increase in the network lifetime. In [15] have introduced a deterministic clustering protocol for energy saving which, as per the claim of the paper, reduces processing element overhead cost to automate the sensor network, which are

getting reflected in the system lifetime. Besides, this research work claims that the approach approximates and accentuates an ideal interpretation for sensible energy consumption in an ordered WSNs. In [16] have incorporated predictive cluster head selection using fuzzy-based scheme for WSNs. This research work introduces a parameter called the rate of recurrent communication in addition to the remaining power of any nodes is used to decide the cluster head. The fuzzy logic method evaluates the Cluster Head Selection Probability which is based on the node's previous communication history to decide the Cluster Head. The rate of recurrent communication of sensor node is found to yield better results compared to the earlier works. The network model and the processing Flow of FBDF-TBC routing protocol is described in Figs. 1 and 2 respectively. Further, this section explains energy consumption model, data compression model and the data fusion model of proposed FBDF-TBC. The following assumptions hold good for our proposed architecture.

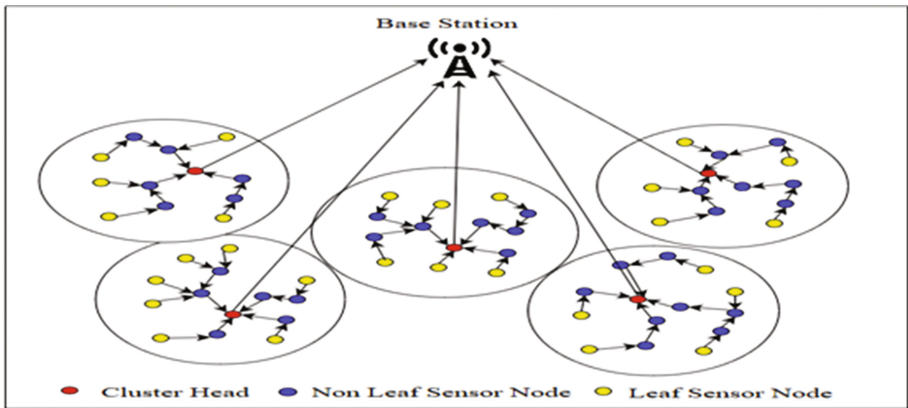


Fig. 1. FBDF-TBC network model

1.3 The Proposed Architectural

The network consists of N nodes equally distributed in a square sensing area and the base station (BS) is far away from the environment that is being sensed.

- BS has unlimited energy resource. The initial battery powers of all the sensor nodes are same initially and the batteries are not rechargeable.
- After the network is deployed, all the sensor nodes and the BS are stationary.
- Each sensor node has same processing and sensing capabilities.

1.4 The Energy Consumption Model

The sensor node majorly comprises of four modules: a power unit; a processing element; a sensing unit; radio frequency transmission unit which consists of an amplifier,

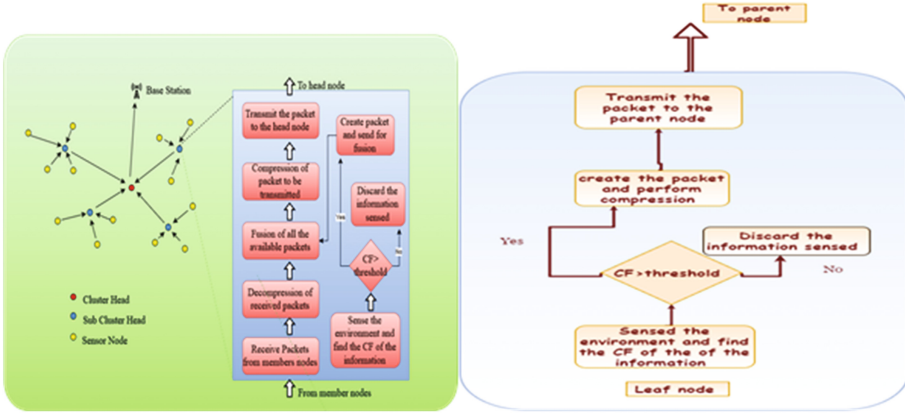


Fig. 2. Process Flow diagram of FBDF-TBC

antenna, and receiver/transmitter circuits. Since the primary objective of this research work is to develop an energy efficient sensor routing protocol that provides accurate sensing information, the energy for transmission and reception is also considered, the energy required to perform data fusion is also taken into consideration. To calculate the transmission energy, the following equations are considered.

$$E_T(q, d) = \begin{cases} q * E_{elec} + q * E_{fs} * d^2, & d < d_{co} \\ q * E_{elec} + q * E_{mp} * d^4, & d > d_{co} \end{cases} \quad (1)$$

Where E_{elec} is the energy consumed by the electrical circuits, q is the size of the packet in bits, d is the space between any twofold nodes, d_{co} is the crossover distance, E_{fs} and E_{mp} are the energies consumed by the amplifiers for distances shorter than d_{co} and distances larger than d_{co} respectively. For receiving a packet of q -bits, the energy consumed is.

$$E_R(q) = q * E_{elec} \quad (2)$$

Hence for a parent node, the energy consumption for a single round

$$E_{parent}(i) = n * E_R(q) + E_F(q) + E_t(q, d(i, j)) + E_S(q) + E_G(q) \quad (3)$$

Where ‘ n ’ represents number of children that have transmitted the packets to the parent node, $E_F(q)$ is the energy for performing data fusion, $E_S(q)$ and $E_G(q)$ are the energies of sensing and generating packets respectively. For the cluster head nodes, the energy consumption is same as $E_{parent}(i)$ except for that $d(i, j)$ is replaced by $d(i, BS)$ where BS is the base station. For the leaf node, the consumption of energy depends on whether the node transmits the data or not. The transmission of the packet is decided by the Type-2 fuzzy logic system. If the sensed data is of greater confidence, then the packet is generated and need to be transmitted otherwise the sensed data is discarded. But for parent nodes, even though the data sensed by it is of lower confidence, it has to

still perform fusion on the packets obtained from its child nodes if any. Thus the energy consumption of a child node is.

$$E_{\text{child}}(\mathbf{j}) = E_t(\mathbf{q}, \mathbf{d}(\mathbf{i}, \mathbf{j})) + E_S(\mathbf{q}) + E_G(\mathbf{q}), \text{ if transmission takes place} \quad (4)$$

$$E_{\text{child}}(\mathbf{j}) = E_S(\mathbf{q}), \text{ if packet is not transmitted} \quad (5)$$

1.4.1 The Compression Model

In WSNs, compression method used to decrease the energy depletion of a node to transfer a packet. By performing compression on the packet to be transmitted, its size is reduced considerably thereby reducing the amount of energy needed to transmit it. However, the most crucial part is to choose an efficient method of compression because the processing capabilities are limited at sensor nodes. This research work opted Distributed Source Coding (DSC) to perform a lossless compression of correlated data values from various sensor nodes.

1.5 The Proposed Routing Methodology

Originally, the inspiration for the development of FBDF-TBC protocol is derived from the extensive analysis of LEACH-C routing algorithm, energy efficient PEGASIS and Type-2 fuzzy logic. The FBDF-TBC routing protocol comprises of the following stages, (a) Clustering and cluster head selection, (b) Cluster Tree formation, (c) Data fusion, (d) Data transmission in the network.

1.5.1 Clustering and Cluster Head Selection

Initially, the configuration (location) and residual energy details are known by BS by transmitting and receiving messages between live nodes and BS. BS separates the entire WSN nodes into five clusters centered on their corresponding proximity. After the cluster creation, a cluster head is selected for each cluster by the BS. Basically, the cluster head alone has the capability to communicate with the BS. Additionally, all the nodes in the cluster transmit to cluster head through a tree based cluster. For example, consider the scenario of 100 live sensor nodes distributed in an area of $100 \times 100 \text{ m}^2$ (refer to Fig. 3). The node at the location (50,175) is the BS and the sensor nodes are distributed into five clusters. The cluster heads are marked with squares and all the other nodes are the non-CH sensor nodes. Further, the member nodes of every cluster are arranged as a minimum spanning tree.

In every iteration, the CH for each cluster is selected by the BS. A sensor node is selected as CH by considering the following two parameters: (a) energy remaining in the sensor node and (b) distance of the node from the base station. The energy E_{average} is used to decide the CH. The node that has residual energy higher E_{average} and paramount cost function is selected as CH for that cluster.

E_{average} Can be calculated as follows:

$$E_{\text{average}} = \frac{\sum_{i=1}^{n\text{Alive}} E_{\text{residual}}(i)}{n\text{Alive}} \quad (6)$$

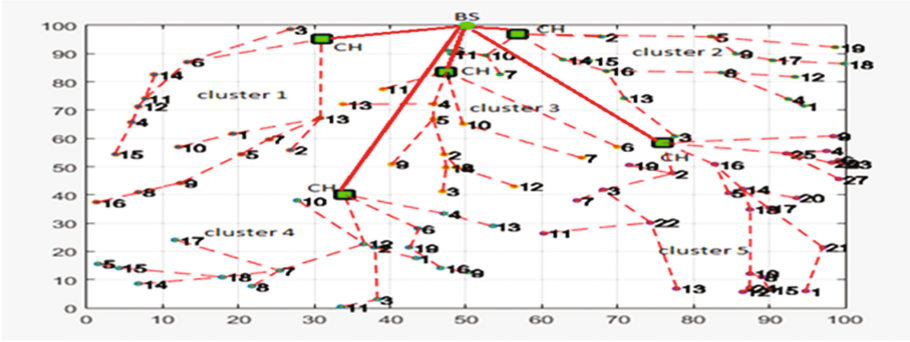


Fig. 3. 100 nodes in a 100 × 100 m² area divided into clusters and a minimum spanning tree is constructed in a round

Where ‘nAlive’ represents the number of live nodes in the cluster and $E_{residual(i)}$ is the remaining energy of the i th node in the cluster. The cost function of a sensor node in a cluster can be calculated as follows:

$$cost(i) = \frac{w_e}{w_d} \times \frac{E_{residual(i)}}{d(i, BS)} \tag{7}$$

Where $d(I, BS)$ is the Euclidean distance between the sensor node and the base station, w_e and w_d are the cost factors of residual energy and distance respectively. The cost factors need to be set appropriately. A Greater w_e value means the available energy of the important node while selecting CH and vice versa.

Algorithm 1. Area Division and Cluster Head Selection

1. INPUT: n-no of cluster required, BS-Base Station node, r- no of round, List of alive nodes
2. OUTPUT: Tree with nodes connected in MST
3. for every node in {Alive Nodes List} doing
 - a. Broadcast NAME packet that contains its ID, residual energy, and Location to BS
4. end for
5. // for each round
6. for x to 1 to r
 - a. Partition the whole network into 'n' clusters with same sub size network
 7. for every cluster(sub-network) in the network do
 - a. Compute the mean energy and Choose the cluster head that the maximum cost function
 8. end for
 9. go to Algorithm 2 // Cluster Tree formation phase (MST)//

1.5.2 The Cluster Tree Formation

In this phase, the sensor nodes belonging to respective cluster that are constructed into a minimum spanning tree using Prim's algorithm in such a way that each tree has a minimum sum of weights.

Algorithm 2. Cluster Tree Formation

```

1. weight[CH]=0
2. weight[other nodes]=MAXIMUM
3. {TREE [BS]}=BS
4. while { Alive Nodes List}!= 0} do
    a. Find out a node i in { Alive Nodes List }, where Weight
       [node i] is minimum
    b. last=node i
       for every node j in { Alive Nodes List } do
        i. if (node j != last) then
            1. if (Weight [node j] > d (node j, last)) then
                a. weight [node j] = d (node j, last)
                b. {TREE [node j]}=last
            2. end if
        ii. end if
    c. end for
    d. Remove the last in { Alive Nodes List }
5. end while
6. generate a TDMA schedule for all the member nodes in every
   cluster
7. send the TDMA schedule and the details of minimum tree based
   clustering in network
return {TREE}

```

2 The Sensor Data Fusion

In any WSN, the sensor nodes transmit the sensed information in the form of packets. However, it is not advisable to transmit all the packets received by a node since there may be uncertainties and redundancies in the sensed data. The redundancy in packets leads to unnecessary energy consumption and bandwidth wastage. Also, the sensors may produce some erroneous information due to many reasons like evolving environmental conditions and manufacturing defects. If this erroneous information are transmitted to BS, it will seriously affect the outcome of the decisions made by the WSN. This research work employs Type-2 fuzzy logic to prevent the redundant data from getting transmitted to BS. After the completion of CH selection and MST construction, every CH generates a TDMA schedule and disseminate it to every member of cluster to deliver data. Each live sensor node in the network is associated with a Type-2 Fuzzy Logic Controller (FLC). The FLC finds the confidence factor (CF) of the data sensed by the sensor based on the current sensor condition. Thus each sensor node

generates packets that consist of both the data and the confidence factor of that data. There are three types of sensor input data; Temperature, Humidity, Signal to Noise Ratio (SNR). For each input data, the expected value and its uncertainty are represented by the covariance and mean matrix. The values are then normalized to a value between [0, 1]. The inputs to the fuzzy system are distributed into three levels: Low, Medium, and High. The output of the FLC is the consequent which is broadly divided into five levels: Very Low, Low, Medium, High, and Very High. As there are three states in each input variable (Low, Medium, and High) and there are three variables (Temperature, Humidity, Signal to noise ratio) there is a total possibility of $3 \times 3 \times 3 = 27$ inference rules. The inference rules are shown in Table 1.

Table 1. Inference rules

Rule no.	Input variable			Output
	Temperature	Humidity rate	Signal to noise	Confidence factor
1	High	High	High	Very high
2	High	High	Medium	Very high
3	High	High	Low	High
4	High	Medium	High	High
5	High	Medium	Medium	High
6	High	Medium	Low	Medium
7	High	Low	High	Medium
8	High	Low	Medium	Medium
9	High	Low	Low	Low
10	Medium	High	High	Medium
11	Medium	High	Medium	Low
12	Medium	High	Low	Low
13	Medium	Medium	High	Medium
14	Medium	Medium	Medium	Medium
15	Medium	Medium	Low	Low
16	Medium	Low	High	Medium
17	Medium	Low	Medium	Low
18	Medium	Low	Low	Very low
19	Low	High	High	High
20	Low	High	Medium	Medium
21	Low	High	Low	Low
22	Low	Medium	High	Medium
23	Low	Medium	Medium	Low
24	Low	Medium	Low	Very low
25	Low	Low	High	Low
26	Low	Low	Medium	Very low
27	Low	Low	Low	Very low

The determination of whether the values of the sensor nodes are in the conventional range is performed by the FLC. If the value is in the accepted range the output of FLC is 100%. If the value is out of range then the FLC generates the CF for data collected. The confidence factor is $0\% \leq CF_n \leq 100\%$. Each sensor node compares the confidence factor of the data sensed against a threshold value or cut-off value. This cut-off is set by the users to determine whether the fuzzy amount produced should be measured or not. If the confidence factor of the data is fewer than the cut-off assessment then the data sensed is discarded. Else, the data is transmitted to the parent node. The confidence factor is calculated for the information that has been sensed by the parent node which is used to decide whether to use the information for fusion or to discard it. If the packets need to be discarded then packets received from its child nodes are fused. Otherwise, the data sensed by the parent node is also fused with the data of its children nodes and the fused data is sent to its parent node. The fusion performed by all the non-child nodes is as follows:

$$FD = \frac{(CF_1 \times D_1) + (CF_2 \times D_2) + (CF_3 \times D_3) + \dots + (CF_n \times D_n)}{CF_1 + CF_2 + CF_3 + \dots + CF_n} \quad (8)$$

Where D_1, D_2, \dots, D_n the data are received by the parent node from its child nodes of one kind and CF_1, CF_2, \dots, CF_n are the confidence factors of the corresponding data and FD is the fused data that will be transmitted to its parent node. As data from different nodes are fused together into FD and also that the data that is used for fusion is of high confidence value the FD is robust and is of higher certainty. The FD is calculated independently for each type of sensor nodes and hence we have a set of FDs instead of a single FD. The set of FDs is represented as a vector V_{FD} .

$$V_{FD} = \{FD_1, FD_2, FD_3, \dots, FD_m\} \quad (9)$$

Where m is the number of different types of data being sensed

Suppose a parent node has three child nodes that sense temperature and the three temperatures are 30 °C, 25 °C and 20 °C their corresponding confidence factors are 0.50, 0.75 and 0.65 respectively. Then the FD for the temperature is found to be 25.60. Similarly, the FDs are calculated for other data as well and the vector consisting of FD will be $V_{FD} = \{25.60, 53.2, 38\}$. Then the consequent of the new data is found. This vector is then passed by the parent node to its parent only if the consequent of the new data is changed. But if there is a modification in the arrangement it does not mean a correct exposure. It is instead measured a likely occurrence in the region that is being monitored. The BS regularly processes the received data to determine whether it is an event or not.

Algorithm 3: Data Fusion and Data Transmission

If (node I in {TREE} is parent node) then

- a. If (the data packet need decompressing) then
 - i. For each 10 bits in compressed data packet do
 1. Y =high 7 bits as side information
 2. S_x = lower 3 bits
 3. Calculate $S_y=H*Y^T$
 4. Calculate $X=[S_x \text{ xor } S_y]^T$
 5. Search X in Z_x which has $Z_x \text{ xor } Y = X$
 6. Append Y and X into data packet decompression
 - ii. End For
 - b. End If

//the received data packets contain $D1, D2, D3, \dots, Dn$ data and corresponding confidence factors $CF1, CF2, CF3, \dots, CFn$ //

//The parent node has the input data T -Node Temperature, H -Humidity Ratio, N -Signal to Noise Ratio //
 - c. $CFp = FLC(T, F, N)$
 - d. If $CFp \leq \delta$
 - i. Discard the data collected
 - ii. FDn =Received Data Packet from each child node will be fused
 - iii. Consequent= $FLC(FD1, FD2, FD3 \dots FDn)/n$ is number of child nodes for
Else
 - iv. Generate packet for the collected data
 - v. $FDn+1$ = Received Data Packets from each child node will be fused along with the packet generated for the collected data
 - vi. Consequent= $FLC(FD1, FD2, FD3 \dots FDn, FDn+1)$
 - e. End If
 - f. If the resultant was not altered
 - i. Ignores the conventional data
 - ii. Send only the nodes data
 - g. Else
 - i. Send the Consequent to parent node
 - ii. IF the event sensed by parent node
 1. Report the event
 - iii. End IF
 - h. End If
2. End If
3. If (the node is child node)
- a. $CFc = FLC(T, F, N)$
 - b. If $CFc \geq \delta$
 - i. Data packet data and CFc will be sent to its parent node
 - c. Else
 - i. Discard the data
 - d. End If
4. If (the data packet need compressing) then
- a. For each 14 bits in the data packet do
 - i. Y =high 7 bits as side information
 - ii. X = lower 3 bits
 - iii. Calculate $S_x=H*Y^T$
 - iv. Append Y and S_x^T into new packet
 - b. End For
5. End If

3 Simulation Parameters

To validate the performance of the projected algorithm, the FB-DFTBC, DFTBC, and LEACH-C are implemented using MATLAB where the simulation parameters are given in Table 2.

Table 2. Parameters for simulation

Description of parameter	Value
1. Area of simulation	100 × 100 square meters
2. Number of nodes in network	100
3. E_{elec} (radio electronic energy)	50 nJ/bit
4. E_{fs} (radio free space)	100 pJ/bit/m ²
5. E_{mp} (radio amplifier energy)	0.013 pJ/bit/m ⁴
6. E_{init} (initial energy of node)	2 J
7. Packet size	500 bytes
8. Base station location	50,175
9. Channel type	Wireless channel
10. Number of clusters	5
11. Simulation time	3600 s

4 Evaluation and Results

4.1 The Simulation Results

Firstly, Fig. 4 showcases that the proposed approach enhances the overall network lifetime than their counterparts. Here, the X-axis represents the lifetime of the network in terms of a number of rounds whereas, Y-axis symbolizes the number of nodes alive. From Fig. 4, it is obvious that the FBDF-TBC (blue in color) have more rounds or longer network lifetime compared to LEACH-C and DFTBC, respectively. Besides, Fig. 4 also describes the comparison between the FBDF-TBC with and without compression (red in color). In short, it enhances the overall network lifetime of WSN from a minimum of 16% (compared to FBDF-TBC without DSC) to a maximum of 76% (compared to LEACH). After that, Fig. 5 represents the time of the death of the last node in the network when the location of Base Station is varied.

It is clear from the graph that the number of rounds in DFTBC and FBDF-TBC relatively decreases when the BS is moved farther, but it remained stable with LEACH-C. Subsequently, Fig. 6 quantifies the percentage of dead nodes for five different routing protocols: FB-DFTBC, DFTBC with DSC, DFTBC without DSC, PEGASIS, and LEACH. The X-axis represents the percentage of dead nodes whereas the Y-axis contains the time in terms of a number of rounds at which the particular percentage of nodes are dead. The graph clearly depicts that FB-DFTBC routing protocol takes longer to attain a larger percentage of dead nodes that all other protocols. So it can be stated that it increases the lifetime of the network and the sensor nodes. Packet delay which is defined as the time difference between the time at which the

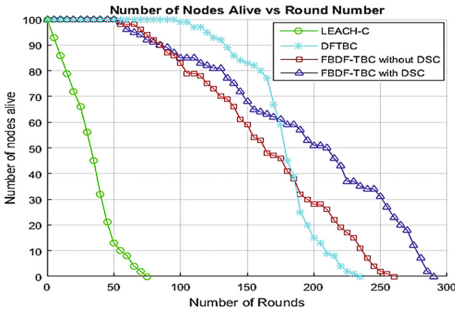


Fig. 4. Number of live nodes with the change in rounds

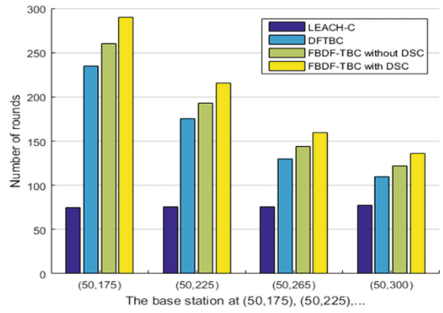


Fig. 5. Death of the last node when the position of BS changes

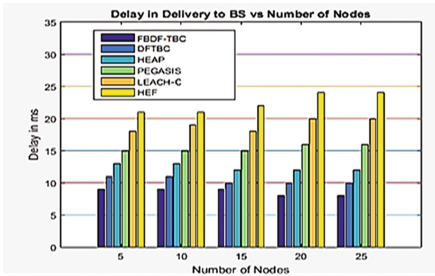


Fig. 7. Delay in the delivery of packets to BS

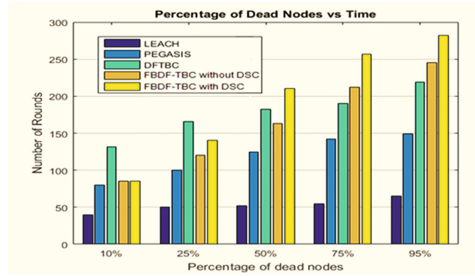


Fig. 6. Percentage of dead nodes when BS at 50,175

packet is created and the time at which it is actually received by the BS is depicted in Fig. 7. It showcases the packet delay of the following WSN routing protocols: FB-DFTBC, DFTBC, HEAP, PEGASIS, LEACH and HEX. Here, the X-axis is the number of nodes that are alive in the sensing environment while the Y-axis is the delay in delivery to BS in milliseconds (ms). It is clear from the graph that the delay is the minimum for FB-DFTBC.

5 Conclusion and Future Enhancements

Generally, energy preservation is the major focus in any wireless sensor network research. With the similar objective, this research also work proposes a new approach for energy-efficient clustering and compression of data packets before actually sending it to the base station. Subsequently, it is also proved from the experimental results that the proposed algorithm greatly minimizes the energy consumption of the network which in turn improves the overall network lifetime. Further, it also enhances the sensing accuracy by eliminating the data redundancy. In concise, the simulation results clearly

concludes that the energy efficiency of the proposed algorithm is higher than its peers thereby improving the network lifetime from a minimum of 16% to a maximum of 76%. The proposed approach can be further enhanced by incorporating sub-clustering as well as fault tolerant characteristics which are our ongoing research work.

References

1. Kumar, N., Tyagi, S., Deng, D.: LA-EEHSC: learning automata-based energy efficient heterogeneous selective clustering for wireless sensor networks. *J. Netw. Comput. Appl.* **46**, 264–279 (2014)
2. Yu, J., Feng, L., Jia, L., Gu, X., Yu, D.: A local energy consumption prediction-based clustering protocol for wireless sensor networks. *Sensors (Switzerland)* **12**, 23017–23040 (2014)
3. Tan, N.D., Viet, N.D.: DFTBC: Data Fusion And Tree-Based Clustering Routing Protocol for Energy-Efficient in Wireless Sensor Networks, vol. 326, pp. 61–77. Springer, Cham (2015)
4. Sengaliappan, M., Marimuthu, A.: Improved general self-organized tree-based routing protocol for wireless sensor network. *J. Theor. Appl. Inf. Technol.* **1**, 100–107 (2014)
5. Liu, Z., Yang, X.: An application model of fuzzy clustering analysis and decision tree algorithms in building web mining. *Int. J. Digit. Content Technol. Appl.* **23**, 492–500 (2012)
6. Mammu, A., Hernandez, S.K., Jayo, U., Sainz, N., de la Iglesia, I.: Cross-layer cluster-based energy-efficient protocol for wireless sensor networks. *Sensors (Switzerland)* **4**, 8314–8336 (2015)
7. Su, S., Yu, H., Wu, Z.: An efficient multi-objective evolutionary algorithm for energy-aware QoS routing in wireless sensor network. *Int. J. Sens. Netw.* **13**(4), 208–218 (2013)
8. Izadi, D., Abawajy, J.H., Ghanavati, S., Herawan, T.: A data fusion method in wireless sensor networks. *Sensors (Switzerland)* **2**, 2964–2979 (2015)
9. Zhang, Z., Liu, T., Zhang, W.: Novel paradigm for constructing masses in Dempster-Shafer evidence theory for wireless sensor network's multisource data fusion. *Sensors (Switzerland)* **4**, 7049–7065 (2014)
10. Chen, J., Han, X.: The distributed source coding method research based on clustering wireless sensor networks. *Int. J. Sens. Netw.* **4**, 224–228 (2015)
11. Bagci, H., Yazici, A.: An energy aware fuzzy approach to unequal clustering in wireless sensor networks. *Appl. Soft Comput. J.* **4**, 1741–1749 (2013)
12. Meghanathan, N.: Stability-based and energy-efficient distributed data gathering algorithms for wireless mobile sensor networks. *Ad Hoc Netw.* **19**, 111–131 (2014)
13. Heinzelman, W., Chandrakasan, A., Balakrishnan, H.: An application specific protocol architecture for wireless microsensor networks. *Proc. IEEE Wirel. Netw.* **4**, 660–670 (2002)
14. Gupta, I., Riordan, D., Sampalli, S.: Cluster-head election using fuzzy logic for wireless sensor networks. In: *Annual Conference on Communication Networks Services*, pp. 255–260 (2005)
15. Aderohunmu, F., Deng, J., Purvis, M.: A deterministic energy-efficient clustering protocol for wireless sensor networks. In: *Proceeding of the Seventh International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*, pp. 341–346 (2011)
16. Natarajan, H., Selvaraj, S.: A fuzzy based predictive cluster head selection scheme for wireless sensor networks. In: *International Conference on Sensing Technology*, pp. 560–566 (2014)