Research On Energy Efficient Hierarchical Clustering-Based Leach Protocol for Wireless Sensor Networks

Izzatilla Khidoyatov
Tashkent University of Information Technologies named after Muhammad al-Khwarizmi, Uzbekistan

Doston Hasanov
Tashkent University of Information Technologies named after Muhammad al-Khwarizmi, Uzbekistan

Azizbek Temirov
Tashkent University of Information Technologies named after Muhammad al-Khwarizmi, Uzbekistan

Temirbek Toshtemirov
Tashkent University of Information Technologies named after Muhammad al-Khwarizmi, Uzbekistan

Ravshanbek Khidirov
Tashkent University of Information Technologies named after Muhammad al-Khwarizmi, Uzbekistan

Received: 28 Feb 2026 | Received Revised Version: 23 Mar 2026 | Accepted: 29 Apr 2026 | Published: 18 May 2026

Volume 08 Issue 05 2026 | Crossref DOI: 10.37547/tajet/Volume08Issue05-09

Abstract

The aim is to reduce the energy consumption of nodes and maximise the lifetime of the sensor network. The original LEACH (Low Energy Adaptive Clustering Hierarchy) protocol selects the head cluster based on random number generation and does not consider the residual energy of the nodes. In this paper, we propose a modified head cluster selection algorithm and cluster formation algorithm. Cluster formation is based on the nearest neighbour distance algorithm, which is an improvement over the k-means algorithm for forming the clustering structure. The simulation results show that the proposed protocol has improved the occurrence of nodes in a non-working state, the percentage of alive nodes, the residual energy, and the throughput of the sensor network.

Keywords: Network uptime, clustering, wireless sensor network, residual energy.

Cite This Article: Izzatilla Khidoyatov, Doston Hasanov, Azizbek Temirov, Temirbek Toshtemirov, & Ravshanbek Khidirov. (2026). Research On Energy Efficient Hierarchical Clustering-Based Leach Protocol for Wireless Sensor Networks. The American Journal of Engineering and Technology, 8(05), 73–82. https://doi.org/10.37547/tajet/Volume08Issue05-09

1. Introduction

Hierarchical clustering is an important approach to reduce energy consumption and improve data transmission efficiency in wireless sensor networks. This approach divides sensor nodes into clusters and provides control through cluster heads. Therefore, Table 1 analyzes the main features, advantages, disadvantages, and application possibilities of hierarchical clustering protocols, based on scientific evidence.

Table 1. Comparative analysis of hierarchical clustering protocols in wireless sensor networks

Protocol name	Description	Advantage	Disadvantage	Application area	Evaluation
LEACH	Divides sensor nodes into clusters and changes the cluster head in each round.	Reduces energy consumption, selects cluster heads in turns, and extends the network lifetime.	Cluster-head selection has a high level of randomness.	Small and medium-scale wireless sensor networks.	Most suitable
PEGASIS	Connects nodes in a chain structure and transmits data sequentially.	Provides high energy efficiency.	May cause high transmission delay.	Long-term monitoring systems.	Good
PEGASIS	Connects nodes in a chain structure and transmits data sequentially.	Provides high energy efficiency.	May cause high transmission delay.	Long-term monitoring systems.	Good
TEEN	Transmits data only when a predefined threshold value is exceeded.	Detects emergency situations quickly.	Not convenient for continuous monitoring.	Temperature, fire, and danger monitoring.	Suitable for special cases
APTEEN	Combines periodic and event-based data transmission based on TEEN.	Flexible; supports both real-time and periodic monitoring.	Has a more complex structure.	Complex monitoring systems.	Good
HEED	Selects the cluster head based on residual energy and communication cost.	Provides good energy balancing.	The computation process is more complex than LEACH.	Large-scale sensor networks.	Very good
Other protocols	Use various optimization methods.	Adapt to specific problems.	Not universal.	Special-purpose sensor systems.	Moderate

In wireless sensor networks, hierarchical clustering protocols are important for reducing energy consumption, increasing data transmission efficiency, and extending the network lifetime. The LEACH, PEGASIS, TEEN, APTEEN, and HEED protocols discussed in the table serve the same general purpose, namely, effectively organizing the activities of sensor nodes. Among these protocols, LEACH stands out as the most suitable protocol. Because it has a simple structure, balances energy consumption by periodically replacing cluster heads, and is convenient for practical application in wireless sensor networks. Especially in small and medium-sized monitoring systems, the LEACH protocol provides an effective solution in terms of energy efficiency, reliability, and simplicity.

Sensor network nodes have very limited battery power, memory capacity, and radio frequency [1]. Saving the energy of nodes is very important. Different clusters are formed using different algorithms for energy consumption, such as k-means, k-Sep, and the nearest neighbor distance algorithm [2]. After the cluster is formed, a head cluster is selected in each cluster based on different parameters. There are two main approaches to wireless sensor networks. The least energy transfer and direct transfer methods may not provide a good trade-off between the energy of the nodes and the load [3]. In direct transfer, the sensor nodes send their data directly to the sink; as a result, there is a single connection between the base station and the head cluster. The least energy transfer is used in the way of minimum cost [4]. To solve this type of problem, LEACH is proposed, which guarantees that dynamically created clusters have an easily distributed energy load. In Leach, the head cluster is selected based on probabilistic methods, generating a random bit between 0 and 1 [5]. Cluster formation is considered an operation that the sensor node decides to perform.

They must connect the head clusters among different alternatives. In the proposed algorithm, additional factors such as residual energy, throughput, and node rank are used to decide the head cluster selection [6]. For example, if an advertisement message is received by a sensor node X from three specific head clusters P, Q, and R. The main parameters for selecting the head cluster include distances, if the distance parameter Q is selected as the head cluster, but this is a bad choice because it has the lowest residual energy of all [6].
There are two phases in the basic leach: cluster setup phase and data transfer phase. The group setup phase is a probabilistic approach used in the LEACH [2] protocol for group selection. We select an extreme value for each group node. An arbitrary value between 0 and 1 is generated and we compare this arbitrary number with the extreme value. If the value is less than the extreme value, the node is selected as the group leader for the current cycle, otherwise it acts as a group member. In the second phase of the protocol, the group leader performs preprocessing and merging of the data. The master cluster sends the data to the base station. The LEACH-C algorithm is an improved version of the LEACH algorithm [7, 9]. In LEACH-C, we use a data-driven routing algorithm. The data-driven routing algorithm finds the routing information while transmitting the message between the sensor nodes, and then the routing information is updated. Usually, the data-driven algorithm allows for directed diffusion and SPIN, and TL-LEACH is an improved algorithm. This algorithm has two master clusters, a primary master cluster and an auxiliary master cluster.

The location protocol is used when the WSN data is not available for each node. When a GPS module is available in the network, PEACH works for location. PEACH creates a cluster without joining and without scheduling messages. It provides randomization-based clustering algorithms to organize sensor nodes into clusters in wireless sensor networks. The most useful option is used to calculate the master cluster.

LITERATURE REVIEW

The nodes are placed in different clusters and each cluster has its own master cluster. In this regard, there are various clustering methods such as adaptive clustering, hierarchical clustering and distribution clustering. Master clusters are selected based on various parameters such as minimum Euclidean distance, highest available energy among sensor nodes and maximum node level. In each cluster, all cluster member nodes send data to the master cluster. There are two main approaches for data transmission, which are experimented as direct transmission and minimum transmission energy. In indirect transmission, data is transmitted immediately from the master cluster to the external base station, which consumes more energy due to the master cluster, in the minimum transmission energy method, the node in the sleeping state that is not in the working state is quickly woken up due to being far from the collecting node, so this method reduces the transmission time and re-detection of the master cluster. In the main transmission, there are two phases, namely, cluster establishment phase and data transmission phase. In LEACH [2], during the cluster establishment phase, cluster heads are selected based on a probabilistic approach. When selecting an extreme value for nodes, an arbitrary value between 0 and 1 is generated and an arbitrary number is compared with the extreme value. If the value is smaller than the extreme value, the node is selected as the head cluster for the current cycle, otherwise it acts as a member of the set. The second stage of the protocol is the preprocessing and merging of the head cluster data. The head cluster sends the processed data to the base station and There are two main approaches for cluster formation, known as k-means and k-NN. In the k-means algorithm, n nodes are divided into k clusters and the Euclidean distance between the nodes is calculated. The starting point is the center of each cluster and the algorithm is run for several cycles. After each cycle, new clusters and centers of the cluster are generated and finally the center point of the cluster is determined. In k-NN clustering, all nodes that are closer are examined based on the Euclidean expansion, and initially node 1 is assumed to belong to cluster 1. On this basis, the distance from the second node to the first node is estimated, and if this separation is greater than the extreme value, it is added to the new cluster. This process is completed in less time and requires fewer iterations than the k-means algorithm. The extreme value is calculated as follows [6]:

$$T(n) = \begin{cases} \frac{P}{1-p\left(r \bmod \frac{1}{p}\right)} & \text{if } n \epsilon G \\ 0 & \text{else} \end{cases} \quad (1)$$

where p denotes the probability of the beginning of the set, the current cycle value is r, and G is the group of centers.

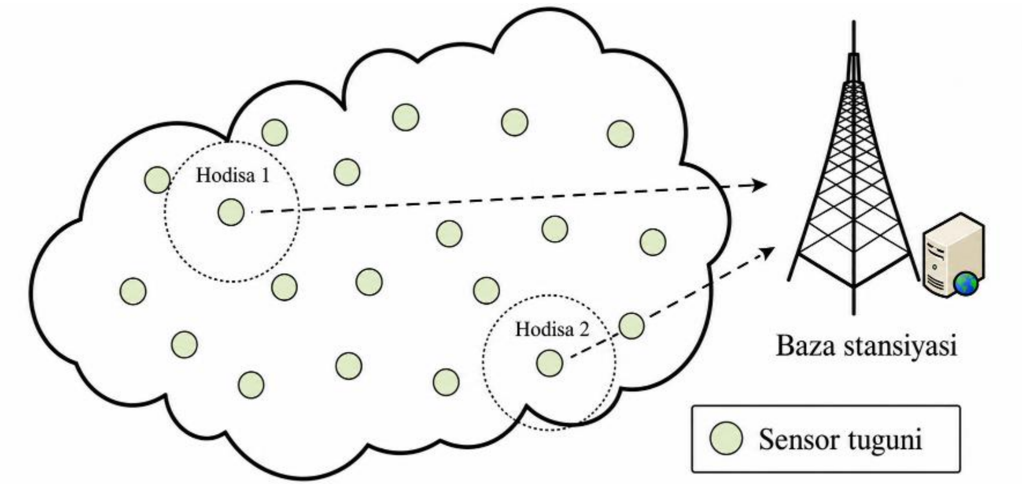

Fig.1. Wireless sensor network model

P is the head cluster at the end of the iterations, and it sends an announcement message to its group member. Each unassembled head node selects the collector head cluster for this cycle, based on the signal quality received from each collector head, which requires the least consistency, the imperative. Each node determines which group it belongs to, and learns to connect to the group head node by receiving a connection request message. The group head node establishes a time division multiple access to the group head center and communicates this plan to all member centers within the group.

ENERGY EFFICIENT ROUTING ALGORITHM BASED ON CLUSTERING IN WIRELESS SENSOR NETWORKS

In developing a new scheme to reduce energy consumption used in cluster formation and head node selection algorithm, the head cluster selection technique is based on the highest ratio of residual energy and distance from the base station. The cluster formation technique is based on the threshold distance and the threshold value is defined as the "average distance of all data points to their center of distribution". If X is the center of all data points, the threshold is defined as the average distance of all data points to X. In the k-means clustering algorithm, which is also known as partition-based clustering, the sensor nodes are grouped into different clusters. This algorithm is based on the Euclidean distance. The location of different sensor nodes and their clustering are shown in the graph below.

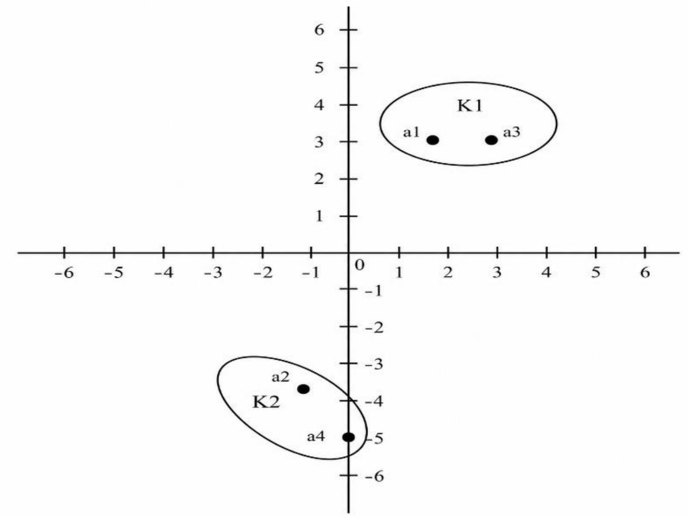

Figure 2. Cluster formation of nodes

Modified cluster formation approach. The modified cluster formation approach calculates the distances between sensor nodes and determines the distance matrix, and with the modified approach, a threshold distance is set and its value is determined by the average distance of all data points to the center of the distribution. x is the center of all data points, the threshold is set based on x as the average distance of each data point. Initially, a node is assumed to belong to the first cluster, and the distance of the second node from the first node is estimated. If the distance is greater than the threshold distance value, then the second node is added to the new cluster. If the distance is less than the threshold distance, then the second node is added to the previous cluster. This process is continued to consider the remaining nodes to form a cluster. The clusters are determined by calculating the distances of the remaining nodes from the nodes, comparing the minimum value from the threshold, and deciding to join the existing cluster or a new cluster member one by one.

Algorithm for calculating the boundary distance

1. Read the coordinates of the nodes (x_i, y_i)

2. For (i, j = 1 to n) {

Calculating distances between nodes using the Euclidean distance formula

$$d_{ij} = \sqrt{(x_i - x_j)^2 - (y_i - y_j)^2}$$

3. Find the mean value (M) of the nodes (x_i, y_i)

For (i= 1 to n) {

{

$$M_x = \sum_1^n \frac{x_i}{n}$$

$$M_y = \sum_1^n \frac{y_i}{n}$$

M = (M_x, M_y)

M(M_x, M_y) Calculates the distance of nodes from x_i, y_i

D_{Mi}= distance ((M_x, M_y) (x_i , y_i)

Calculate the average value D_{Mi}

$$M_D = \sum_{i=1}^{n} \frac{D_{Mi}}{n}$$

}

Set M_D otherwise

Border =M_D

Modified cluster formation algorithm

1. Consider the first node in the first cluster
2. For (i = 2 to n){
3. The remaining node (n_i) to (n_1 *to* n_{i-1}) distance to
4. Minimum distanceD_{min} = { d_1, d_2, $d_3, \dots \dots . d_{i-1}$}
5. S= install D_{min}
6. If (D_{min}< D_{min}){
7. (n_i) knot taken (x_i, y_i) if added to the node D_{min}, (n_i) The node is added to that cluster.

   }

   Else {
8. Node (n_i) joins the new cluster

   }

}

The current approach to selecting a primary cluster is based on generating random numbers between zero and one, and setting the threshold value according to the following equation.

$$T(n) = \begin{cases} \dfrac{P}{1-p\left(r \bmod \frac{1}{p}\right)} & \textbf{if } n \in G \\ \mathbf{0} & \textbf{otherwise} \end{cases} \quad \textbf{(2)}$$

where p is the probability of the optimal head cluster, n is the total number of nodes, r is the last round, and G is the group of nodes, indicating that they did not select the head cluster in the last 1/p iterations.

This algorithm is illustrated by the following example. Let there be four sensor nodes located at different positions, a_1 (3, 3), a_2 (-1, -4), a_3 (2, 3), a_4 (0, -5). The distance matrix is calculated using the Euclidean distance formula. The Euclidean distance between two nodes with coordinates (x_1, y_1) and (x_2, y_2) is calculated using the following formula.

$$\mathbf{d_{ij}} = \sqrt{(x_i - x_j)^2 - (y_i - y_j)^2} \quad (4)$$

2-jadval. Sensor tugunlarining masofa matritsasi

Tugun	a_1	a_2	a_3	a_4
a_1	0	8.1	1	8.5
a_2	8.1	0	7.6	1.41
a_3	1	7.6	0	8.2
a_4	8.5	1.41	8.2	0

In this approach, the center of all data points within the boundary distance (x)={(3+(-1)+2+0)/4,(3+(-4)+3+(-5)/4)} = {1, -3}

Distance from a1 to x (d1) = 6.32

Distance from a2 to x (d2) = 2.23

Distance from a3 to x (d3) = 6.08

Distance from a4 to x (d4) = 2.23

The average of these distances = {(d1+d2+d3+d4)/4} = 4.21 = The marginal distance value.

Step 1: a1 is placed in its own cluster 1, where K1 = {a1}

Step 2: If d (a1, a2) = 8.1, otherwise its value a2 is placed in another cluster K2 = {a2}.

Step 3: Calculate the distance between a1 and a2 in a3 and calculate the distance matrix of a3 and the threshold value d(a3, a1) = 1< and this is the closest to a1 and place a3 in cluster 1, so K1 = {a1, a3}

Step 4: Calculate the distance from a4 to a1, a2, a3 using the distance matrix of a4 and the nearest to a2 using d(a4, a2) = 1<4.21 and K2 = {a2, a4}.

where the center point of cluster K1 = (2.5, 3) and the center point of cluster K2 is (-.5, -4.5).

Modified approach to head cluster selection. In the modified approach, the head cluster is selected based on the highest ratio of energy and space remaining from the base station. The existing approach selects the head cluster based on random number generation. The existing approach does not consider the residual energy in head cluster selection. In the modified head cluster selection algorithm, the distance parameters from the base station are taken as the residual energy of the node in the cluster.

The cluster formation of four sensor nodes a1 (3, 3), a2 (-1, -4), a3 (2, 3), a4 (0, -5) located at different positions is shown below, where all sensor nodes have the energy amount of the distance of each sensor node from the base station in the cluster and there are two cluster sets K1 and K2, and the head cluster should be selected in K1 and K2. The following table (residual energy and distance ratio) is constructed using the residual energy equation and the Euclidean distance formula.

Table 3 Residual energy and distance ratio

Node	**Distance from base station (3, 5) (d_i)**	**Residual energy in July E_i**	**Ratio ($\frac{Ei}{di}$)**
a1 (2, 3)	2	35	.175
a2 (-1, -4)	9.84	.25	.0254

a3 (3, 3)	2.23	.40	.179
a4 (0, -5)	10.44	.30	.0287

The energy dissipation for transmitting an L-bit message from the transmitter to the receiver over a distance d is defined as:

$$E_{Tx}(l,d) = \begin{cases} L.Eelec + L.\in fs.d^2 & if\ d < do \\ L.Eelec + L.\in mp.d^4 & if\ d \geq do \end{cases} \quad (1)$$

where E_{Tx} is the energy dissipated by the transmitter and Elect is the energy dissipated per bit to operate the receiver-transmitter circuit. The parameters ∈fs and ∈mp depend on the transmitter amplifier used in the energy model. The amplifier parameter for the free-space propagation model is ∈fs. The amplifier parameter for the two-beam ground propagation model is ∈mp. The crossover distance, d0, can be obtained from. The Euclidean distance formula is shown below.

$$d = \sqrt{(x_i - x_j)^2 - (y_i - y_j)^2} \quad (5)$$

In cluster K1, node a3 is selected as the head cluster because it has the highest ratio of residual energy to distance from the base station. In cluster K2, node a4 is selected as the head cluster because it has the highest ratio of residual energy to distance from the base station.

Proposed head cluster selection algorithm

1. Primary cluster selection (n)

2. For (i= 1 to n) {

Residual energy and distance from base station for each node d_i considering E_i residual energy E_iCalculate the ratio. R_i and base station distance (d_i) as $R_i = \frac{E_i}{d_i}$ Sorts in ascending order of ratio.

}

The node chooses a potential parent cluster and broadcasts its ratio to the cluster nodes, adding n_i to the new cluster. All cluster nodes receive the ratio of the current parent cluster and compare it with their own ratio.

If (if node ratio > cluster node ratio is obtained) {

The node becomes the cluster head and their ratio to all cluster nodes moves to stage 3.

}

Else{

The current node is the cluster head.

}

RESULTS AND DISCUSSION

The simulation parameters for a cluster-based wireless sensor network with an area size of 100m×100m are shown in Table 1. The number of sensors is n = 200, and these nodes are both simple and complex, and are evenly distributed across the area. As can be seen from Figure 3, the number of nodes that are not in working condition appears after more cycles in the proposed hierarchical clustering-based protocol, and from this perspective, there are more nodes that are not in working condition, which reduces the data transmission rate and increases more traffic between sensor nodes.

Table 3. Simulation parameters

Parameter	Value
Simulation area (m)	100 x 100

Number of nodes	200
Radio propagation model	Ikki nurli yer
Channel type	Simsiz kanal
Optimal probability (P)	0.2
Initial energy	0.9
Base station location	(50,50)
Data packet size	4000 bit
Data collection energy	5nj bitli signal
Maximum number of rounds	100
Free spatial propagation (regular)	10 pj bit / m^2
Multipath propagation (regular)	0,0013 pj bit / m^2
Percentage of advanced nodes (m)	0.1

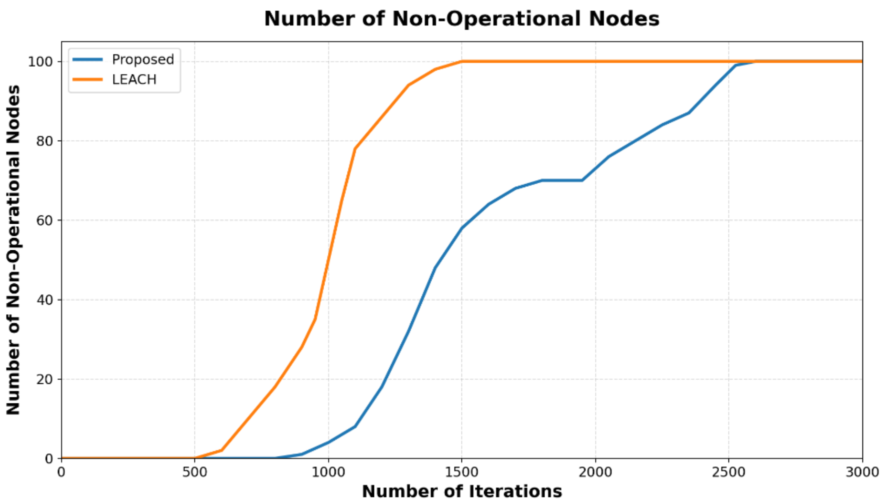

Figure 3. Number of occurrences of nodes that are not in working condition in a particular type

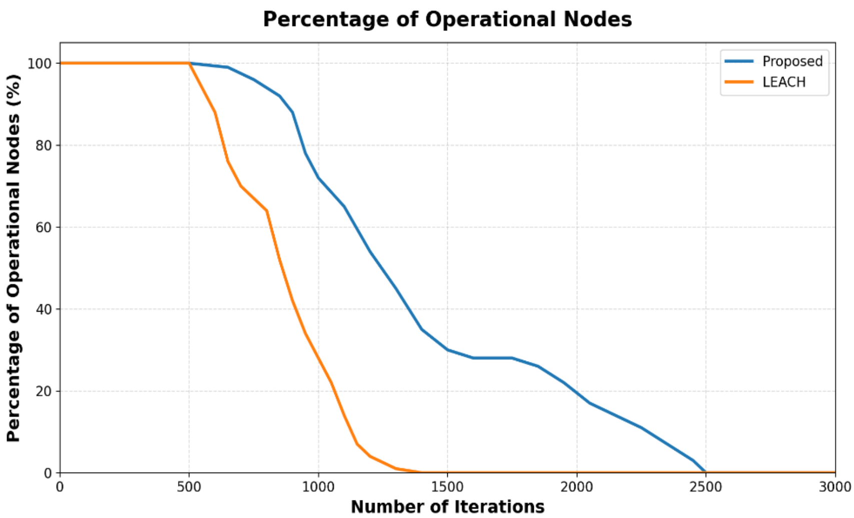

Figure 4. Percentage of nodes in working condition in each type

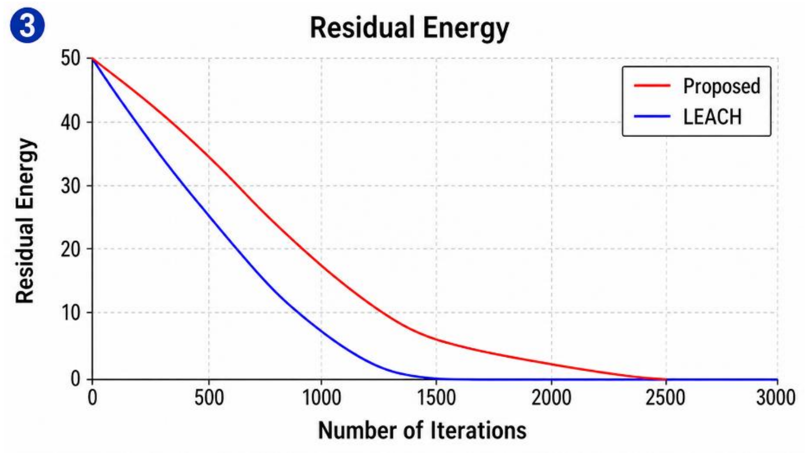

Figure 5. Percentage of residual energy in each type

Figure 6 shows the throughput of the entire network. Throughput is expressed as a ratio of message length to transmission time. In the proposed protocol, the transmission time to send data from the transmitter to the receiver is less, so it increases system performance and consumes less energy.

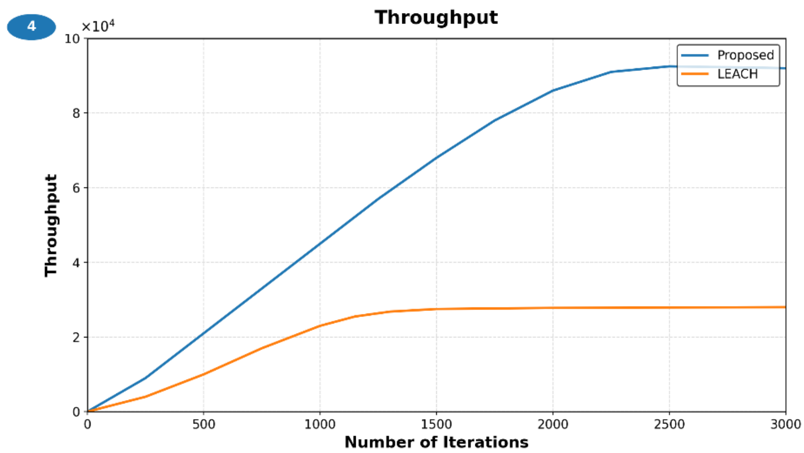

Figure 7. Maximum throughput of nodes in the current iteration

CONCLUSION

In order to improve energy efficiency in wireless sensor networks, a modified cluster head selection algorithm and an improved cluster formation algorithm were proposed. In the proposed LEACH protocol, the cluster formation process is organized based on the Euclidean distance and threshold distance criteria. Here, the threshold distance is determined by the average distance of all sensor nodes to the distribution center. This approach allows the cluster formation process to be performed in a smaller number of iterations. In the modified cluster head selection algorithm, the ratio of the residual energy to the distance to the base station for each sensor node in the cluster is calculated. These ratios are sorted in ascending order, and the node with the highest value is selected as the cluster head. Experimental results show that the proposed LEACH protocol is more efficient than the standard LEACH. In particular, a decrease in the number of nodes that are not in working condition, a longer retention of residual energy, an increase in throughput, and a longer retention of the share of nodes in working condition ensured the overall stability of the network. Future research plans to study the performance of the energy-efficient clustering algorithm in heterogeneous networks, and to consider additional parameters such as node level, sensitive area coverage, and data aggregation when selecting the primary cluster.

REFERENCES

1. Jennifer Yick, Biswanath Mukherjee, Dipak Ghosal. "Wireless sensor network survey", Published by Elsevier, 14 April 2008, Page No: 2292 – 2330.
2. S. Bandyopadhyay and E. J. Coyle. An energy efficient hierarchical clustering algorithm for wireless sensor networks. In Proceedings of the 22nd Annual Joint Conference of the IEEE Computer and Communications Societies (Infocom 2003), April 2003
3. S. Bandyopadhyay and E. J. Coyle. Minimizing communication costs in hierarchically clustered networks of wireless sensors. Comput. Networks, 44(1):1–16, January 2004.
4. E. J. Duarte-Melo and M. Liu. Analysis of energy consumption and lifetime of heterogeneous wireless sensor networks. In Proceedings of Global Telecommunications Conference (GLOBECOM 2002), pages 21–25. IEEE, November 2002
5. E. J. Duarte-Melo and M. Liu. "Analysis of energy consumption and lifetime of heterogeneous wireless sensor networks", In Proceedings of Global Telecommunication, conference, IEEE 2012, page No: 21-25..
6. Gopi Saminathan Arumugam and Thirumurugan Ponnuchamy. EE-LEACH: development of energy-efficient LEACH Protocol for data gathering in WSN EURASIP Journal on Wireless Communications and Networking (2015) 2015:76.
7. K. A. Abdul Nazeer, M. P. Sebastian Improving the Accuracy and Efficiency of the k-means Clustering Algorithm; 2009. Hossein Jadidoleslamy. "An Introduction To Various Basic Concepts Of Clustering Techniques On Wireless Sensor Networks", International Journal Of Mobile Network Communications & Telemetric Vol.3, February 2013
8. Jiguo YU , Yingying QI , Guanghui. "An energy-driven unequal clustering protocol for heterogeneous wireless sensor networks", J Control Theory April 2011vol.9 page No: 133–139.
9. Ravneet Kaur1, Deepika Sharma2 and Navdeep Kaur. "Comparative Analysis of Leach and Its Descendant Protocols in Wireless Sensor Network", International Journal ofP2P Network Trends and Technology- Vol.3, 2013, page No: 1-5.
10. D Guo, L Xu. "LEACH Clustering Routing Protocol for WSN. "In Proceedings of the International Conference on Information Engineering and Applications" 2013, page No: 153–160.