RESEARCH ARTICLE

RSP Science Hub

International Research Journal on Advanced Science Hub
2582-4376
Vol. 05, Issue 07 July
www.rspsciencehub.com

Check for updates

http://dx.doi.org/10.47392/irjash.2023.048

An Improved Data Aggregation Method to Minimize The Energy Consumption and Increase Life Time In WSN

M Raju[1], *K P Lochanambal*[2]

[1]*Research Scholar, Department of Computer Science, Government arts College, Udumalpet, India*

[2]*Assistant Professor, Department of Computer Science, Government Arts College, Udumalpet, India*

Email: rajum@skasc.ac.in

Article History

Received: 19 June 2023
Accepted: 21 July 2023
Published: 28 July 2023

Keywords:

Wireless Sensor Networks;
Data Aggregation;
Clustering;
IDAC

Abstract

Many applications exist for WSN, including but not limited to environmental monitoring, exploration, and military surveillance. When data relevant to several applications is gathered by the WSN's nodes and sent to the sink, it may be analysed and repurposed. Sensing data is often sent from the sensor to the sink via multi-hop routing. To identify the network event, the sink stores the data in a database, processes it using control instructions, and then assesses the results. (Wu, Lee, and Chung) Data aggregation reduces the amount of data that must be sent by processing data locally inside the network.This paper suggested clustering method called "Improved Data Accumulation Clustering (IDAC)" emphasises the formation of clusters with the goals of load balancing and lifespan extension for the network. First, the cluster heads are selected depending on their distance from the hub. The base station is located outside the sensing area. At initially, we feed each sensor the same amount of energy. All sensors can simultaneously gather data and it has been send to hub. The BS is aware of the positions of every sensor node, also nodes at top level of cluster have communication capabilities. cluster head probability is denoted by the symbol p. Once a node has served as cluster leader for 1/p rounds, it may serve in that capacity again. Initial node energy levels are represented by the symbol Emax.

1. Introduction

Clustering is a popular method for improving energy efficiency. Clusters are formed by the nodes scattered throughout the sensing region. As the name states CH is responsible for data collecting and transmission for each cluster (Forero, Cano, and Giannakis). Clustering can lower network energy usage by allocating resources and redistributing load. The cluster heads collect data and send it back to the BS directly else indirectly via various cluster heads. Clusters can form using one of two methods:

i. To classify the nodes, a variety of criteria are employed, including their proximity to one another, the size of the cluster, the number of nodes, and their distance from the base station. The nodes in each cluster are classified as cluster heads based on their energy level, transmission range, and closeness to the sink.

ii. Cluster heads are chosen at random and then encourage adjacent nodes to join them.

The cluster heads can be selected via a variety of methods. The cluster heads are possibly the nodes with the most resources. Occasionally, cluster heads may be chosen at random. If a node serves as the cluster leader for an extended period of time, its power may be rapidly depleted. Fixed cluster head clusters may not be proportionate in terms of the number of cluster members, resources utilized, and data sent to the sink. These unequal clusters can also result in complex runtime complications (Jafari and Es-Hagi).

As resources deplete, the same node cannot continue to function as cluster head indefinitely. As a result, cluster head responsibilities are rotated across nodes to extend the network's lifespan. Cluster heads can be replaced based on events or time. In order to balance resource utilization, the network will be reclustered at a particular time or when a predetermined event happens. Cluster heads may be reelected at the local or global levels (Deif and Gadallah). Individual clusters reselect the cluster heads in local reclustering, but all clusters in the network reselect the cluster heads in global reclustering (Sneha and Nagarajan).

2. Contribution of Clustering in Minimizing Energy Utilization in WSN

The areas where Wireless Sensor Networks (WSN) is deployed necessitate constant energy supply. Despite their widespread use, sensors may be difficult or impossible to recharge in certain situations, such as war zones or with wildlife. As a result, companies seek to maximize their available energy utilization. Clustering is critical in lowering energy consumption in WSN. Choosing the cluster head (CH) of a sensor cluster from the best sensor nodes saves energy consumption. The most crucial component of clustering is CH selection, which, depending on the distance at which packets are sent, helps to save energy. Additionally, any cluster-specific issues should not affect the entire network, allowing for more effective usage of the network. The base station, sink and the CH must be installed properly for proper communication in clustered WSNs. Positional differences enable efficient data flow on the network and a longer network lifetime. (Nagarajan and Karthikeyan)

While clustering maximizes power efficiency with wireless transmission, the absence of a system for choosing the best (CH) node which increases the complication of data collection and energy demand of sensor nodes. Along with, the imbalance in energy demand may cause nodes against fail prematurely, thereby reducing the network's lifespan. To maximize energy conscious and network durability of the WSN, the researchers proposed that a technique known as Hybrid Particle Swarm Optimization (HPSO) is coupled with improved low-energy ACH (HPSO-ILEACH) for CH selection in data assembling scenarios. This method calculates the cluster head, the distance among the cluster's nodes, along with remaining liveliness of the nodes using HPSO. ILEACH has been utilized in-order-to reduce liveliness consumption throughout the clustering procedure by modifying the cluster head. Hence HPSO-ILEACH deals for data aggregation and energy conservation has been implemented successfully (Abreu and Mendes) (M and Krishnaveni) (Sasikumar and Khara).

A unique data aggregation method depending on node clustering and the Extreme Learning Machine (ELM) was proposed. Hence suggested method filters sensor data at each sensor node using the Kalman filter (KF). The similarity of the node data and the cluster head data, as experience-based by cosine similarity and bulk information, at home in cluster the mote, and ELM is involved to the data in the cluster head to enhance data aggregation performance. ELM was found as an approximation and feed forward neural network (NN) that is effective for regression classification and clustering. This does not require any iterative processes, which significantly minimizes the computing time required during the training phase. (Istwal and Verma) However, with ELM, the hidden layer's input weights and biases are created at random, leading in a loss of prediction stability and model precision. Furthermore, this doesn't enable multifaceted data analysis, it is essential for recognizing fairly complex data. The Mahalanobis distance-based redial basis function (MDRBF) is added to ELM projection moment to accurately initialize it's model's weights and biases and accelerate the tuning phase. The proposed strategy may effectively reduce the loss function while improving clustering accuracy. Computer simulation shows that the suggested technique surpasses the real ELM and existing clustering schemes in period of clustering precision, network longevity and energy efficiency (Manasi and Joshi)

The cluster head's energy and the distances between CH and the BS are very changeable in the

LEACH protocol due to the arbitrary nature of cluster formation.

Cluster leaders are responsible for gathering and integrating data from shared nodes within their clusters in addition to giving data to the base station. Data transmission costs more energy than data fusion during data gathering and transfer. Cluster heads will shut down rapidly owing to high energy consumption if their current energy is insufficient or if they are too far away from the base station. To address these difficulties, this research presents a novel, upgraded algorithm for balancing the energy burdens of these cluster heads. LEACH-TLCH (LEACH Protocol with Two Levels Cluster Head) is a more advanced version of LEACH that uses the same cluster-head selection and cluster formation techniques as LEACH. If the present state energy of the cluster head is less than the average energy, or if the distance between the cluster head and the base station is greater than the average distance, the common node in this cluster with the most energy appointed as the secondary cluster head (M and V). When a cluster has a secondary CH, that will be in charge of collecting and transferring data from the member nodes to the cluster head, while the CH is only in charge of sending data to the base station. Then CH is responsible for assembling data from member nodes and transferring it to the BS once the data has been fused in a cluster without a secondary cluster head (Osamy, El-Sawy, and Salim). All clusters are separated into two groups: those with secondary cluster heads disseminate status information to the other regular nodes, build a schedule using a TDMA access channel with each node assigned a specific time slot), and send the schedule to the other nodes. When a normal node joins a cluster without the presence of a secondary cluster head, the cluster leaders split the sending time slot among the other nodes. Once each node has gotten its sending time allocation, the stable phase begins (I Khalaf and &abdulsahib)

The Enhanced Hierarchical Multipath Routing (EHMR) protocol chose the next hop node along with highest residual energy and the fewest hops. As a result, a path with minimal latency is built. Every node chosen for the primary along with secondary paths identifies itself as a cluster head. When a node receives a CH announcement message, it joins the cluster head and forms as a group. Data distribution is balanced over both the paths to minimize overloading of one path. This results, load-balanced multiple paths assure dependability, while hierarchical clustering ensures scalability. According to simulation results, the proposed EHMR protocol may provide high data rates as well as low latency paths. (M, V, et al.)

An innovative, lightweight, and energy-efficient function-based data aggregation approach for a hierarchy of cluster-based WSN is used to create energy-efficient data aggregation for cluster-based wireless sensor networks. In order to increase the exactness of data aggregation, the exponential moving average (EMA) is used at the node level for data aggregation, and a threshold-based technique is used in-order to identify any outliers. This is to deliver extremely accurate aggregated data to the BS at the CH level, we updated Euclidean distance function, the experimental findings demonstrate that our method delivers highly refined, fused data to the base station while reducing the transmission cost, communication cost, and energy consumption at the nodes and CH's. To save computing resources, this method reduces the number of data comparisons at the CH level. The cluster head receives the partially aggregated data flows from the sensor nodes, here they were further aggregated. Only the data streams of neighbouring nodes are compared for the sake of energy conservation and extending network lifetime. This is due to the relative higher likelihood of spatial and temporal connection among nearby nodes. In doing so, the number of comparisons is reduced to a larger extent, conserving both processing and transmission energy (Zhou and Lin) (Ezhilarasi and Krishnaveni)

3. Improved Data Aggregation Clustering to Improve Energy Efficiency

The "Improved Data Aggregation Clustering (IDAC)" proposed clustering technique concentrates on creating clusters to balance the network load and extend the network lifetime. Cluster heads are initially identified based on their proximity to the base station. Outside of the sensing region is where the base station is situated. At first, the same quantity of energy is given to each sensor. Data can be collected by all sensors and transmitted to the BS. Among all of the sensor node's locations are known to the BS. Thus the data can be sent and received by the cluster heads. p stands for the likelihood of becoming a cluster head. After 1/p rounds, a node that has once become cluster head is eligible to do so again. The energy of each node initially is denoted E_{max}.

Initially the base station gathers the details of the location of the nodes. As the sensor nodes are deployed deterministically using TriCentroid algorithm, the location remains fixed throughout the lifetime of the network.

The sensing region is divided into zones and a cluster head is assigned for each zone. As all the nodes in the sensing region are initially supplied with same amount of energy E_{max}, the cluster heads are elected with probability p where p= 1/number of nodes in a zone. Initially all the nodes in the zone have equal probability and hence the cluster head is selected randomly.

Once the cluster heads are elected, they broadcast their identification to all the nodes of the network. The hop table is maintained in each node, the nodes find the closest cluster head and join them to form actual clusters. The cluster head prepares the data sending schedule and then sends that to their members within the cluster. Only 50% of the nodes in a cluster are in active mode and the remaining 50% nodes in a cluster are in sleep mode. The sleep and active modes are switched on fixed time slots. The CH receives data from each node, compresses the data and sends it to the BS. The new-member CH is elected based on the residual energy of it's member nodes. A threshold is determined in this case. The node with the maximum residual energy becomes the next CH. Cluster heads continue to function as CHs when their energy level falls below a predefined threshold. If the energy that remains exceeds the threshold, the CH asks the cluster members for the residual energy and distance to the BS. The member node along with it's most remaining energy and closest proximity to the BS is chosen as the next CH. This step will be repeated till all of the nodes in the cluster have been depleted. This equation is used for calculating the energy threshold:

$$E_{TS} = E_{init}(CH_i) - 40\%(E_{init}(CH_i))$$

E_{TS} is the energy threshold; E_{init} is the initial energy of the Cluster head in each cluster. The energy threshold is computed each time a new cluster head is selected. This procedure ensures that the energy of the cluster head does not completely deplete and that it can be replenished before the other nodes take over as cluster heads, thereby extending the network's lifecycle and decreasing the number of worn-out nodes.

Hence the residual energy of all nodes fall below the threshold, then the next CH will be determine based on the criteria given below:

1. A node has not come under the category to become as a CH for the past (1/p) – 1 rounds.
2. The Current energy ratio of the node to the initial energy of the node must be equal to one. $\frac{E_{res}}{E_{max}} = 1$
3. The next CH must be nearest to the existing CH where the hop count must be 1.

Data aggregation is performed in the CH's and the BS. In the cluster heads data aggregation is performed in two cycles. In the first cycle the aggregation is performed at the CH's for the nodes which are active during the current time slot and in the second cycle aggregation is performed for the nodes that have become active in the next time slot. The data sensed by the members are sent to the CHs and the CH splits them into multiple datasets. A standard threshold value is set initially to accept and discard the data sensed. Euclidean distance is calculated among the datasets and compared with the standard threshold value. Based on the conditions given below the datasets may be accepted and discarded. Let d_{si} and d_{sj} be the two data sets. The standard threshold S is defined initially. The Euclidean distance between these data sets are calculated using the equations given below:

$$\varepsilon_d = \sqrt{\sum_{1}^{t_s-1} (ds_i n - ds_j n)^2 \leq S}$$

$$\varepsilon_d = \sqrt{\sum_{1}^{t_s-1} (ds_i n - ds_j n)^2 > S}$$

If ε_d is between ds_i and $ds_j \leq$ S, then elements in both sets are similar. In that case the elements in one set are retained and the other is discarded. Also ε_d is between ds_i and $ds_j >$ S, then elements in both sets are not similar. In that case the elements in both the sets will be retained. This process is done by all the cluster heads. After aggregation is completed the refined data is sent to the BS. Then in the

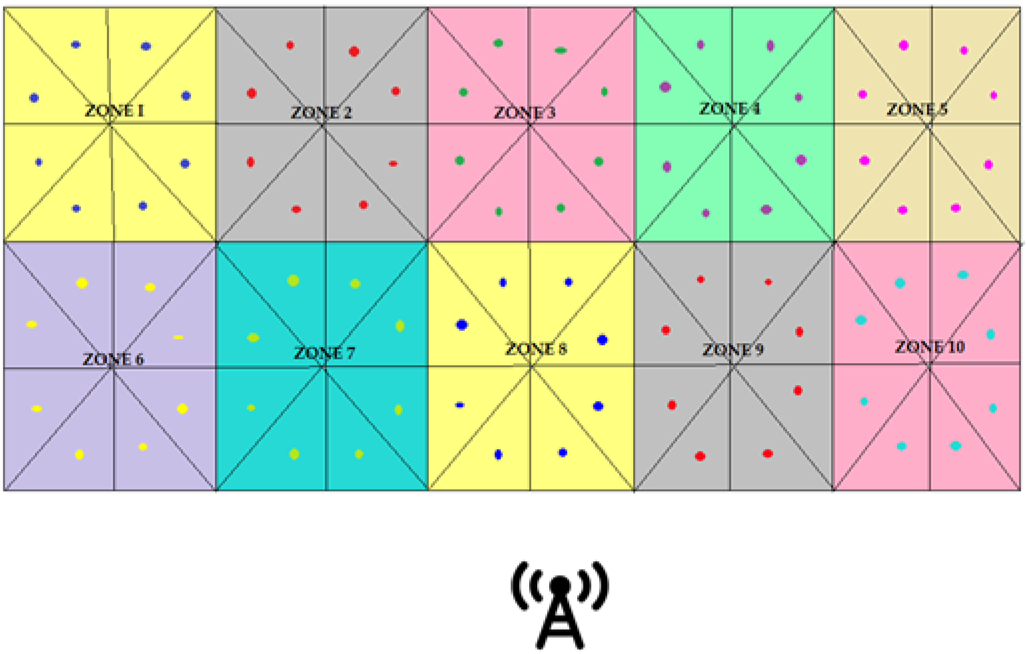

FIGURE 1. **Division of zones using IDAC**

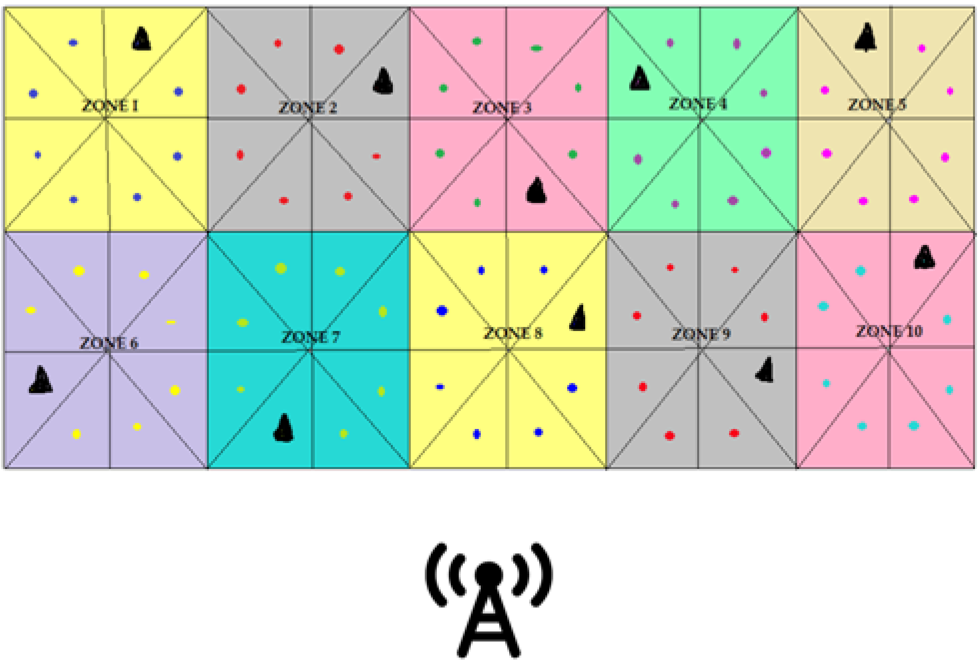

FIGURE 2. **Identifying Cluster Heads using IDAC**

base station the second level of data aggregation is performed similar to the one performed in the CHs. The data received from different CHs are split into data sets by the BS and the Euclidean distance is calculated among the data sets based the conditions given. The data is refined by this level of aggregation and duplicate data is completely removed thereby enhancing accuracy in the sensed data.

4. Results and Discussion

The suggested "Improved Data Aggregation Clustering (IDAC)" technique is intended to build clusters that consume less energy and last longer in networks. In order to demonstrate the efficacy of the proposed algorithm, the parameters cluster size, time required to form clusters, energy consumption during clustering and data aggregation, packet delivery ratio, packet loss ratio, and

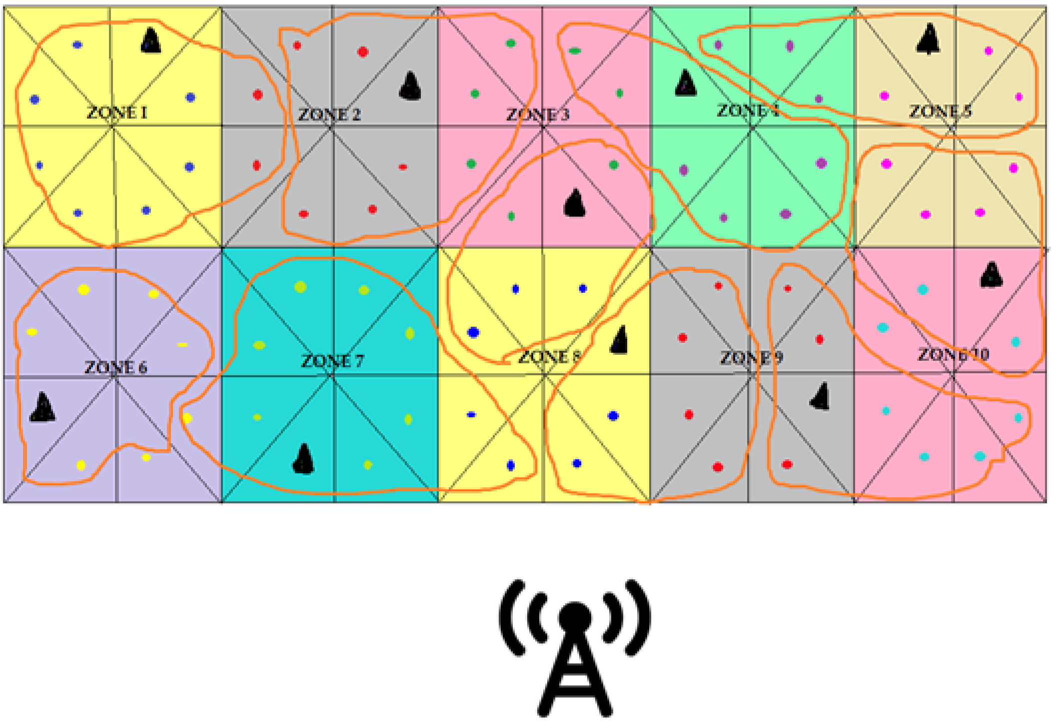

FIGURE 3. **Formation of Clusters using IDAC**

average latency during data aggregation are evaluated and discussed. To analyze the performance of the proposed approach, three current clustering techniques are used: Enhanced Hybrid Multipath Routing (EHMR), Low Energy Adaptive Clustering Hierarchy with Two Level Cluster Heads (LEACH-TLCH), and Energy Efficient Data Aggregation Clustering (EEDA). Table 1 lists the parameters of simulation.

TABLE 1. **Simulation parameters**

Parameter	Value
Network size	300 x 300
No. of sensor nodes	500
Radio propagation range	400 m
Channel capacity	2 M bits/s
Initial energy	1000 J
Hello packets	1600 bits
Simulation time	180 s
ε_fs	10 pJ/bit/m2
ε_mp	0.0013 pJ/bit/m4

4.1. Simulation parameters

The clusters can be balanced or unbalanced. The number of minimum nodes in a cluster refers to the size of a cluster. Our proposed work generates partially balanced clusters as the no. of nodes in each cluster may not be equal. It is called partially balanced as there is no vast difference in the number of nodes in each cluster.

Figure 4 compares the proposed algorithm IDAC to other existing studies and shows that IDAC produces partially balanced clusters. Balanced clusters extend the network's longevity by balancing the energy level of each node inside cluster. By forming clusters of varying sizes, the existing operations make it difficult to manage network traffic and minimize node failures.

Figure4 depicts the time required by the algorithms to generate clusters, and it is clear that the proposed IDAC approach generates clusters faster than the other techniques given in Table 2.

TABLE 2. **Time taken togenerate clusters**

Clustering Approaches	Time Duration (in secs)
LEACH-TLCH	0.27
EHMR	0.15
EEDAC	0.23
IDAC	0.08

Packet Delivery Ratio (PDR) the number of packets delivered to the receiver to the number of packets sent by the sender. Thus proposed algorithm

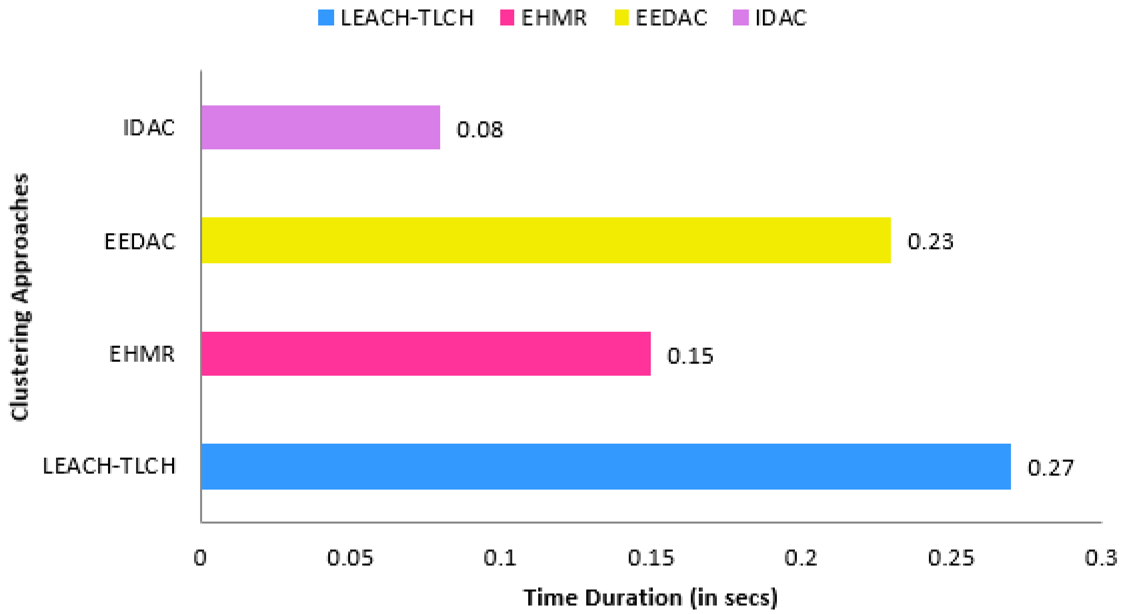

FIGURE 4. Time taken to generate clusters

packet delivery is calculated based on the data packets sent by the cluster heads to the member nodes and vice versa. Each round 100 data packets are transmitted by each cluster head during reclustering and around 500 data packets are transmitted by the member nodes to the cluster heads during data aggregation. Simulation is done by transmitting 200 data packets from the CH to the members. Table 3 shows the details of number of packets delivered after each round.

TABLE 3. Number of packets delivered after each round

Roun	LEACH-TLCH	EHM	EEDAC	IDAC
1	162	179	181	193
2	167	173	184	195
3	159	160	172	189
4	152	156	168	187
5	149	147	157	188

The proposed algorithm shows a constant packet delivery ratio. Table 2 shows the number of packets delivered to the receiver end to the total no. of packets sent. Thus the Figure 5 shows that packet delivery rate is higher in the proposed work compared to the existing works. This assures that the

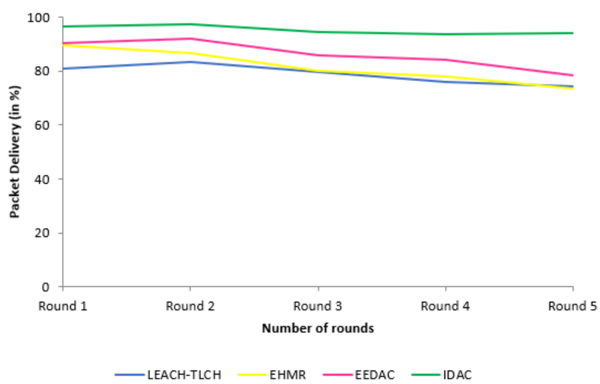

FIGURE 5. PDR of the proposed algorithm with the existing algorithms

proposed algorithm is reliable in terms of data transmission. As the proposed algorithm delivers more number of packets to the destination, it is evident that the packet loss will also be minimal which is depicted in Figure 6.

The time taken to perform data aggregation is given in Table 4 and it is visible that the proposed algorithm consumes less amount of time to aggregate data in comparison to the existing algorithms as shown in Figure 7.

Figure 5.6 shows the consolidation of the time taken to perform clustering and data aggregation where the proposed algorithm out performs the existing ones. The details of consolidation are given

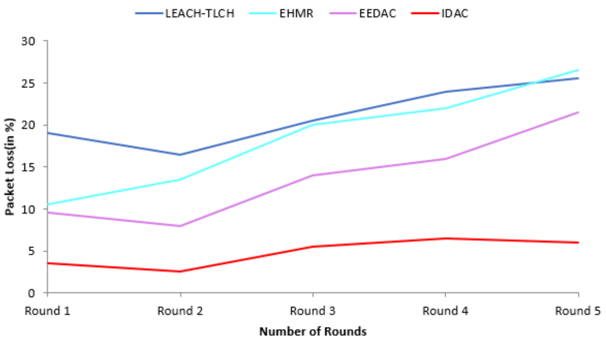

FIGURE 6. **PLR of the proposed algorithm with the existing algorithms**

TABLE 4. **Time taken to aggregate data**

Clustering Approaches	Time Duration (in secs)
LEACH-TLCH	3.46
EHMR	4.15
EEDAC	1.23
IDAC	1.07

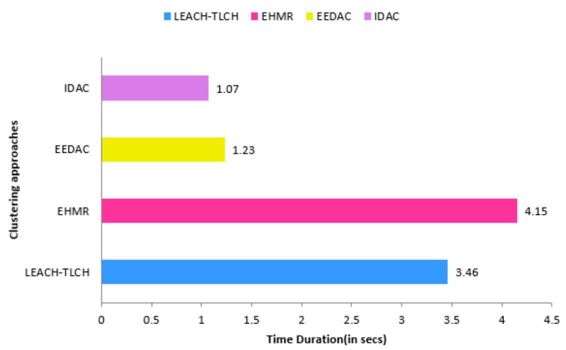

FIGURE 7. **Time consumed for Data aggregation**

in Table 5.

TABLE 5. **Consolidated Time taken to generate clusters and aggregate data**

Clustering Approaches	Time Duration (in secs)
LEACH-TLCH	3.73
EHMR	4.3
EEDAC	1.46
IDAC	1.15

The energy consumed during clustering and data aggregation is shown in Table 5.6. As LEACH-TLCH and EHMR do not have any specified techniques to aggregate data, the time taken by them to

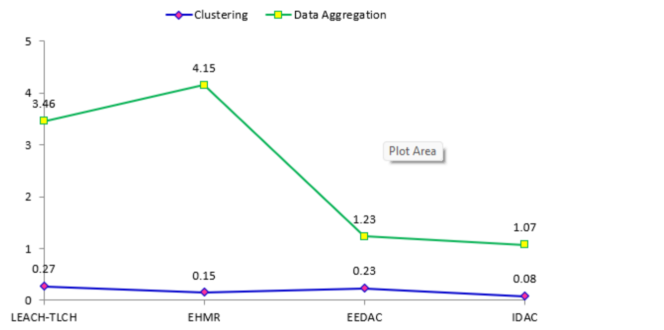

FIGURE 8. **Comparison of Average Delay of the proposed algorithm and existing algorithms with respect to Clustering and Data Aggregation**

aggregate data is comparatively more than the other two approaches. The performance of the proposed algorithm is displayed in Figure 7.

TABLE 6. **Energy Consumption of the algorithmswith respect to Clustering and Data Aggregation**

	LEACH-TLCH	EHM	EEDAC	IDAC
Clustering	112 J	101 J	96 J	87 J
Data aggregation	221 J	231 J	153 J	136 J

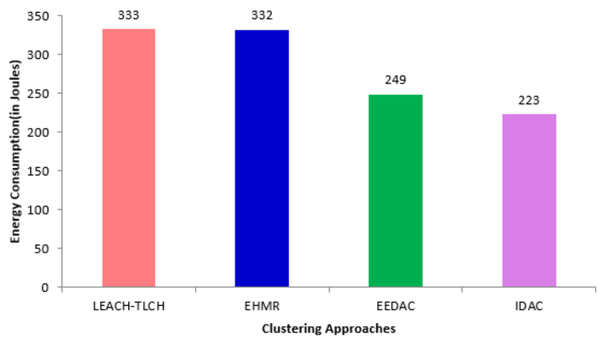

FIGURE 9. **Total Energy Level Consumption of the proposed algorithm compared with the existing algorithms with respect to Clustering and Data Aggregation**

The results from Figure 7 shows that the proposed algorithm consumes less energy than the existing algorithms which can help to prolong network lifetime.

5. Conclusion

The **"Improved Data Aggregation Clustering (IDAC)"** clustering and data aggregation technique is efficient in terms of packet delivery and energy usage. The technique generates clusters of balanced size, which balances the network's burden and lowers the risk of nodes dying suddenly and losing connection. The proposed technology appears to offer reliable communication links that ensure packet delivery while increasing network throughput.

References

Abreu, Carlos and Paulo Mendes. "Deployment of wireless sensor networks for biomedical applications". *2013 IEEE 15th International Conference on e-Health Networking, Applications and Services (Healthcom 2013)* (2013).

Deif, Dina S and Yasser Gadallah. "Classification of Wireless Sensor Networks Deployment Techniques". *IEEE Communications Surveys & Tutorials* 16.2 (2014): 834–855.

Ezhilarasi, M and V Krishnaveni. "A survey on energy and lifetime in wireless sensor networks". *Comput. Rev. J 1* 1 (2018): 30–51.

Forero, Pedro A, Alfonso Cano, and Georgios B Giannakis. "Distributed Clustering Using Wireless Sensor Networks". *IEEE Journal of Selected Topics in Signal Processing* 5.4 (2011): 707–724.

I Khalaf, O and G M &abdulsahib. "Energy efficient routing and reliable data transmission protocol in WSN". *Int. J. Advance Soft Compu. Appl* 12.3 (2020): 45–53.

Istwal, Yasha and Shashi Kant Verma. "Dual cluster head routing protocol in WSN". *2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT)* (2017).

M, Ezhilarasi and V Krishnaveni. "An evolutionary multipath energy-efficient routing protocol (EMEER) for network lifetime enhancement in wireless sensor networks". *Soft Computing* 23.18 (2019): 8367–8377.

M, Ezhilarasi and Krishnaveni V. "An evolutionary multipath energy-efficient routing protocol (EMEER) for network lifetime enhancement in wireless sensor networks". *Soft Computing* 23.18 (2019): 8367–8377.

M, Ezhilarasi, Krishnaveni V, et al. "A survey on energy and lifetime in wireless sensor networks". *Comput. Rev. J 1* 1 (2018): 30–51.

Manasi, N and Joshi. *Parallel K - Means Algorithm on Distributed Memory Multiprocessors*. Spring, 2003.

Nagarajan, M and S Karthikeyan. "A new approach to increase the life time and efficiency of wireless sensor network". *International Conference on Pattern Recognition, Informatics and Medical Engineering (PRIME-2012)* (2012): 2012–2012.

Osamy, Walid, Ahmed A El-Sawy, and Ahmed Salim. "CSOCA: Chicken Swarm Optimization Based Clustering Algorithm for Wireless Sensor Networks". *IEEE Access* 8 (2020): 60676–60688.

Sasikumar, P and Sibaram Khara. "K-Means Clustering in Wireless Sensor Networks". *2012 Fourth International Conference on Computational Intelligence and Communication Networks* (2012).

Sneha, V and M Nagarajan. "Localization in Wireless Sensor Networks: A Review". *Cybernetics and Information Technologies* 20.4 (2020): 3–26.

Wu, C H, K C Lee, and Y C Chung. "A Delaunay triangulation based method for wireless sensor network deployment". *Computer Communications* (2007): 2744–2752.

Zhou, Jiliang and Ziqiang Lin. "Lightweight load-balanced and authentication scheme for a cluster-based wireless sensor network". *International Journal of Distributed Sensor Networks* 17.2 (2021): 155014772098032–155014772098032.

Embargo period: The article has no embargo period.

To cite this Article: Raju, M, and K P Lochanambal. "**An Improved Data Aggregation Method to Minimize The Energy Consumption and Increase Life Time In WSN**." International Research Journal on Advanced Science Hub 05.07 July (2023): 232–241. http://dx.doi.org/10.47392/irjash.2023.048