A Novel Power-Efficient Data Aggregation Scheme for Cloud-Based Sensor Networks

Abhishek Bajpai, Rajkiya Engineering College, Kannuaj, India*

 https://orcid.org/0000-0003-2815-2092

Shashank Yadav, Rajkiya Engineering College, Kannuaj, India

Naveen Tiwari, Rajkiya Engineering College, Kannuaj, India

Anita Yadav, Harcourt Butler Technical University, India

Mansi Chaurasia, Rajkiya Engineering College, Kannuaj, India

ABSTRACT

Sensor nodes are being deployed everywhere as per the applications and real-time data analysis. A major concern of this implementation is the limited battery power and huge data generation. The data redundancy can also be a cause of battery decay. This scheme spends the energy based on priority. This method also uses a mobile agent for the data collection from the sensor nodes. When it is combined with optimal cluster head along with marking of subtle aggregators, it gives a satisfying performance. This approach is divided into three phases: clustering of sensor nodes then computing PEDAS and finally deploy a mobile agent. The approach of PEDAS measures parameters in an optimized manner which develops an energy-efficient system and only spends the energy at the moment when it is needed the most. The proposed model was simulated and verified using network-simulator 3. Implementation and analysis of the algorithm prove that this research study has improved the lifetime of the entire network and also provided a stable and robust network while comparing it with EEDAC and ATL schemes.

KEYWORDS

Cloud Computing, Data Aggregation, Internet of Things, Sensor Network

1. INTRODUCTION

Wireless sensor networks as the name suggests can be defined as the network of several wireless entities primarily various sensors (Roy & Chandra, 2019). Sensors prove as of great utility for determining real-time physical changes as they can pick up changes in the surroundings and update the data related to it through which further recommendations can be justified. Wireless sensor networks have found massive number of applications in recent times. They are used in area monitoring, habitat monitoring, health monitoring, smart cities (Giliberto et al., 2019) etc. The sensor nodes could be deployed effortlessly in difficult terrains and Warfield for military assistance. Also the cost to setup

wireless sensor network is not much. All these factors have contributed in the rapid technological advancements in the area of wireless sensor networks. The sensor nodes deployed to monitor environment, industrial machinery etc. in real time can produce enormous amount of data.(Arena & Pau, 2020) But the sensors deployed are very fragile and have very less computational power. So energy conservation in such type of networks becomes very crucial. The major problem of the network is the limited battery. Higher battery usage causes less network lifetime which may lead to crashing down the network and hardware. Therefore, to provide the longest uptime of the network, the battery issue is to be settled with utmost priority. Various research studies suggested event-based utilization of the sensor motes (Padmaja & Marutheswar, 2018; Zhang et al., 2020). Sleep and wake as the initial method used to control the power consumption of the motes gradually became less feasible and more complex. The sensor motes are expected to finally transmit aggregated data to the sink node to process. Overhead data transmission causes the battery drainage at a much faster pace depending upon the number of hops. Considering the ultimate destination of the useful transmission as sink node it can be worked upon. Although the sink node is privileged to have power in abundance same cannot be assigned to all sensors as cost-measure. Therefore, conserving data aggregation schemes with minimal transmissions to the sink node can be deployed. Some important aspects of this scheme are non-redundancy and complete data aggregation in data collection. Redundancy can be mitigated by efficient data-aggregation techniques as it directly affects the energy requirements.

The research study of data fusion involves two phases of aggregation one at the local level and the other at the global level known as global aggregation. Localization of sensors also plays a vital role for which GPS can't be a solution considering energy requirements. Therefore, some optional methods of localization are RSSI-based, which relies on the relative signal strength identifier, values calculated from the transceiver, and then interpolating to find coordinates. Advanced localization methods allow sink nodes to localize and map out regular nodes thus helps in grouping and classification of nodes based on distance and energy pattern. Using the above idea, a near-perfect clustering or grouping of sensor nodes with optimal disseminations to sink nodes can be drawn. Clustering schemes are beneficial in the various aspects of the sensor network like scalability, energy optimization, and lifetime (Bongale et al., 2020). It frontiers the communication in a local domain and then lets it reach out to the larger group. A group of sensor nodes is known as a cluster and its working paradigms are calculated and regulated by the cluster head as well as data for fusion and analysis purposes which is eventually transmitted to the sink node. Clustering also helps in inter-cluster routing for mobile nodes as the cluster head acts as a bridge between sensor nodes and the final sink node. Cluster head selection is one of the issues which depends on residual energy, number of neighbours, the distance from the base station. The cluster head selection paradigms have been classified into about six different categories as Event to sink directed clustering (ESDC) wait for an event to trigger which when the report is submitted to the sink, load-balanced clustering schemes which form the cluster based on load incurred and all the heavy computations performed by cluster heads, K-means clustering a Machine Learning approach which also allows a distributive computing, LEACH (Low energy adaptive clustering) scheme in which each node has an equal probability to be selected as the cluster head thus sharing the energy and computations and finally HEED (Hybrid Energy-efficient distributed clustering) based on residual-energy and intra-cluster communication cost. There are some other methods as well such as modified LEACH and Weight-based methods. Motivation behind this research was to achieve high Quality of Services (QoS), which includes minimizing data transmission delays, and conserving energy. This research study is divided into the five-section. The first section introduces the technical terms that are being used in this study. In the second section, the study elaborates on the previous research study. The next section describes our proposed approach and network topology that is being implemented for data analysis. In further section analyze the result with previous research study and conclude this study future suggestions.

2. LITERATURE SURVEY

Previous works on enhancing the network life in the network in wireless sensor networks have explored increasing the durability of the network by utilizing the battery efficiently. Roy et al. (Roy & Chandra, 2019) use EEDAC-WSN, in which different control signals are exchanged in form of metadata packets in a cluster. Key idea is to detect significant changes with the help of metadata and then allow cluster nodes to transmit based on residual energy, previous round data packet numbers, hop-count, and timestamps as a combined parametric value. Kumar et al. (N & A, 2020) uses real-time computation of significant parameters and then adjusting according to ATL (Availability throughput lifetime) scheme backed by other real-time measures like network-lifetime and DAGS (Data aggregation support measure). This scheme relies on waking up the sleeping sensor nodes to transmit before the next duty cycle as they would have high residual energies. Latha et al. (2019) deal with energy-efficient routing in the first phase then perform congestion-less aggregation in the second and third phases through ESR (Energy sustained routing) and finally CADC (Congestion aware data collection) and called trust-assisted energy-efficient aggregation protocol TA-EAA. Anup Kumar et al. (2020) take LEACH as one of the prominent approaches for clustering method and proposed by in which intra-cluster aggregation strategy in various rounds with each round having three basic steps clustering step using LEACH, intra-cluster aggregation, and finally data transmission. Zhang et al. (2020) employ Entropy-driven data aggregation which tries to maximize the network lifetime by resolving the energy-hole problem using energy preserving. Here energy-hole problem has been tackled using the gradient deployment algorithm and uses a tree-based aggregation method with the addition of fuzzy logic. Padmaja et al. (2018) use residual energy levels of the neighbor nodes. Each sensor node transmits a hello message with its ID and residual energy while the receiver node adds the sender in its neighbor list upon receiving those messages and a cluster is formed. Finally, a node with the highest residual energy is made Cluster Head. CH roles are rotated to manage energy consumption. Similar to the previous scheme Al-humidi et al. (2019) have added a feature of redundancy-reduction from close neighbor nodes transmissions sucking more power as well. Therefore, idea is to increase information quality while reducing the number of packets and conserving energy thus called a lightweight data collection scheme. This scheme operates at local and global levels both called local and global aggregators respectively. Yuvaraj et al. (2019) use a time-oriented energy location energy availability data rate (LEAD), where scheduling is done via the polling method. The LEAD weight measure is calculated for all the sensor nodes in the deployed region and selected for data collection by the mobile agent. The salient feature is the scheduling of sensors based on energy and location. El Fissaoui et al. (2019) as an improvement of the previous method uses mobile MA (mobile agent) which moves around the already described path calculated via minimum spanning tree algorithm. Yestemirova et al. (2018) stress maintaining multiple sink nodes to avoid a single point of failure. Although, maintaining energy efficiency is a daunting task. Extending on the Mobile agent idea previously described, Lohani et al. (2016) aim at the right planning of the itinerary and selection of the optimal factors for itinerary construction for the mobile agent. The path is planned based on the node's energy, network life increment, end-to-end delay, and aggregation ratio. Asemani et al. (2015) used a learning automata algorithm that works under three criteria: data aggregation, residual energy, and number of hops to sink. This keeps in touch with all the data defined above and adapts itself dynamically as data varies from node to node. The learning algorithm used here is called the INCASE-LA. Another method is given by Shanmukhi et al. (2015) Based on a new model comb-needle discovery support model which uses event-directed clustering as the basic topology of the model. Krishna et al. (2013) employ a trade-off between energy and communication cost keeping primary concern as battery problem and thus assume sink as an infinite source of energy. Here also clustering takes place for proper data aggregation with local and global aggregators appointed keeping primary consideration as a residual battery, latency, and amount of processing required. Sinha et al. (2013) dictated a multilevel strategy for energy-efficient data aggregation which considers data sensing from the physical environment as a continuous and stochastic procedure. It is bent towards aggregating non-redundant data-keeping minimal

transmissions to the sink node rather than doing through less hopping. Different filtering schemes applied to keep the non-redundant data at bay and enrich the sink node with only the most important information which can cost the network in future. Krishnan S et al. (2016) discussed power sufficiency for a longer period such as using a cumulative acknowledge. Similarly, a survey by Al-karaki et al. (2004) sheds light on different types of power-efficient data gathering algorithms with their paradigms and a categorical analysis to provide an initial line of thought. After original clustering protocols, Swaminathan et al. (2017) brought two clustering schemes together. LEACH and PEGASIS are mixed into a new hybrid approach to work in sync as CCMAR (Cluster chain mobile agent routing) in which cluster has been referred to LEACH while chaining went to PEGASIS and ultimately a mobile agent which will walk through the defined path is found put after both the clustering schemes did their part. Another approach given by Ramachandran et al. (2018) called the Honey-hive based efficient data aggregation in wireless sensor networks. It focuses on designing an autonomous topology considering economic and battery conservation and tries to minimize transmission delay and gather data using honey-hive analogy then perform data aggregation. Manoharan et al. (2016) gave the idea of LPEDAP (Localised power-efficient data aggregation protocol) which tries to maintain the network lifetime at an optimum level. It was implemented with the LMST and RNG Topologies. Using the L-PEDAP, the parent node is elected and the shortest route to the sink is decided. Data transmission is periodically not continuous. Implementation moves forward in three volumes of work, topology construction, route discovery, data aggregation. Shobana et al (2021) presented an approach called CSDMA (Cluster Based Systematic Data Aggregation Methodology) in which aggregation is carried out in three levels. The cluster head is selected based on the ranking given to sensors, the ranking of a node is decided on the basis of energy level of the node and the Euclidian distance between the node and the base station. The approach given is efficient, increases network life time and is energy efficient. On the basis of many parameters, we compared few past work. Comparison given below in table-1.

3. METHODOLOGY

Our proposed approach inclined towards the idea of efficient and useful information rendering data aggregation. The work has been subdivided into three phases; the first phase is the clustering phase where large deployed sensor nodes are clustered in their respective zones by a selection of cluster heads using traditional LEACH protocol, using TDMA. We have enhanced it by merging with CDMA. LEACH-CDMA helps to minimize the interferences between the clusters. Then the second phase deals with finding a global support measure value, which has been discussed in great detail in our approach called PEDAS (Power-efficient data aggregation support) with a micro packet which is embedded along with the keep-alive beacon signals and help us process the global value for each node present in each cluster and ultimately inform the sink where data density and power efficiency are best, helping in designing of the most optimized route. Coming to the third and last phase a mobile agent is deployed to get the value rather than long radio transmissions to completely optimize the

Table 1. Comparison of different schemes based on network parameters

Research Study	Stability of Network	Energy Conservation	Fault Tolerant	Enhance Throughput	Traffic Load	Data Security
Roy et al (2019)	✓	✓		✓	✓	
Kumar et al (2020)		✓			✓	
Padmaja et al (2018)	✓	✓		✓		✓
El Fissaoui et al (2019)		✓	✓		✓	
Shobana et al (2021)		✓		✓	✓	

set-up. The mobile agent has already been fed the routing path using a suitable minimum spanning tree algorithm.

Network Topology: This research study aims to increase network lifetime by mitigating data redundancy. In this study sensor, mote is deployed randomly in the field. These sensors are measuring the quantitative data of the environment and communicate by mesh topology. As per our approach, the sensors are dived into different zones as per the proximity to the cluster head (CH). Cluster heads aggregate the data that is being generated by sensors of their respective zones. CH transmits the aggregated data to the sink for further analysis in the cloud server.

3.1 Phase-1: Clustering of Sensor Nodes

There are t cluster-zones among the deployment of the sensors. There is a set of $Gn = \{S1, S2, \ldots Sn\}$ sensor nodes and the cluster head (CH) would be elected from them where $t << n$. Cluster heads (CH) are assigned using low energy adaptive cluster head algorithm along with the CDMA such that it helps to mitigate interferences among the clusters by a different set of CDMA codes. A node will be selected as cluster head or not depends the desired percentage of CHs in the network and number of times the node was selected as CH in the previous rounds. Each node i choose random number between 0 and 1. If this number is less than a threshold, i becomes CH.

P is percentage of cluster heads desired, r is the current round $G_n^{'}$ is the set of nodes that have not been CH's in the last $1/R$ rounds:

$$Threshold(i) = \begin{cases} \dfrac{P}{1 - P * \left(r \, mod \, \dfrac{1}{R} \right)} & if \; i \in G_n^{'} \\ 0 & otherwise \end{cases}$$

Figure 1. Network Topology for this study
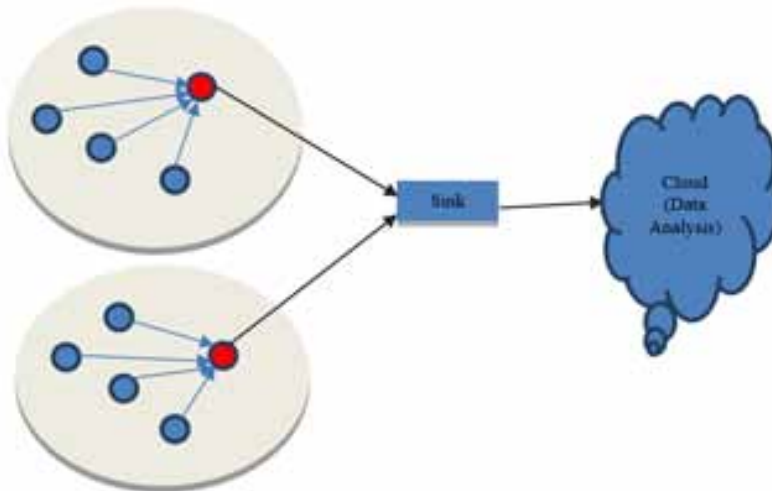
A stochastic algorithm is used to determine the next cluster head in each round. Previously made cluster-heads are not in contention for the next round, therefore the probability of selected as the cluster head is $1/n$.

The most favourable usage of CDMA comes with the benefit that it allows multiple sensor nodes to share the same frequencies and allows the production of different bandwidths on the same power. Then clusters are formed and data transmission to the cluster heads is determined by CDMA channel access, then our approach jumps to the second phase which is a very critical phase as desired calculations and processing are undertaken to identify power-efficient and higher data density regions of nodes.

3.2 Phase-2: PEDAS Processing With Embedded Micro-Packet Mechanism

The second phase deals with the selection of various regions of nodes based on power efficiency and data density. We have defined a global support measure for every node present in the system called PEDAS. Major parameters are the current battery of a node, data density of a node and transmission time to deliver data to its cluster head. PEDAS value for each node is calculated at the sink which further decides which zone or region of nodes to be selected for optimized global battery and information gaining data for the next round. PEDAS value processed at the start of each round at sink node using data related to it from the nodes along with the beacon signal which is used for keep-alive signals by every node to let the sink node know that node requires no maintenance and is ready to move on. Values which are needed to be relayed back are power loss and information gain index value of each node from the previous round. We have also proposed a 3-bit micro packet that can easily be embedded with the already put keep-alive beacon signals without extra overhead. Design and working have been discussed further in detail, but we can relate it to the physical and real aspect of factors that helped us to construct PEDAS measures. The first on the list is an absolute difference in battery life from the previous round in other words power loss for the current round. Our approach tries to optimize global network lifetime by selecting the right number of nodes having a better battery in the current round keeping the network lifetime at a stable state. The second factor is data density reflecting or pointing out nodes that are eager to transmit the data based on their buffer values and the need for its data for the running of the application. If there is such critical data necessary to the system, its information gain index is intensified and selected for the current round at greater priority. Data density is formulated further in the approach determining the amount of data and its information gain index. Last but not least, the factor is the transmission time or time taken to transmit, as it drains the battery. Let's look at the procedure used for relaying the raw materials back to sink in the form of a 3-bit micro packet scheme, where the least significant bit signifies the information gain index and the rest two bits are used to identify the power loss compared to the last round:

$$S_{CH} = \{i: i \in SG_{CH} \ \text{if} \ I_i > 0 \}$$

Micro Packets of size 3 bits will be sent by nodes to the Sink Node comprising data for Information Gain and difference in power from the last beacon sent.

- The 0th bit will represent Boolean value denoting Information Gain:
    - 1 if there is some information gain
    - 0 if there is no information gain

Table 2. Sink Node data

2	1	0

- The other two bits are used to represent the difference in battery percentage from the last recorded value at the sink node.

For this purpose, a **Special Representational Scheme** is used in our methodology. We introduced a Pre-defined Least Significant Bit which is used by the programmer. This **PLSB** is a constant bit, to which the 2 bits representing power difference are added making it a 3-bit pattern.

Example: - Let's assume an incoming Micro Packet = 1 1 1

Using this Special Representational Scheme, we decide when a node becomes ready to transmit the data along with the beacon.

3.2.1 Case-1: If PLSB = 0

If Power Difference is 6, a packet must be sent irrespective of any Information Gain as it is maximum even value denoted by 3 bits.
Else if Information Gain = 1 and Power Difference = 0, 2, 4 then packet must be sent as there is some Information Gain.

3.2.2 Case 2: If PLSB = 1

If Power Difference is 7, a packet must be sent irrespective of any Information Gain as it is an odd value denoted by 3 bits.
Else if Information Gain = 1 and Power Difference = 1, 3, 5 then the packet must be sent as there is some Information Gain.
In this way, we embed two types of information in a 3-bit package which is going to be sent along with the beacon signals used to keep messages alive by the nodes. Now required information to the sink is transmitted, our PEDAS is calculated as follows.
Any sensed values is element of the set $D_i = <d_1, d_2, d_3, ...d_n>$, collected by a particular sensor node S_i at a specific time frame t_j, then PEDAS value describe in Eq. (1).
PEDAS (P) depends on the followings:

1. Battery Life difference of the node from the previous round (d).
2. Data density (D).
3. The time window for which a node transmits data (t).

Figure 2. micro-packet representation
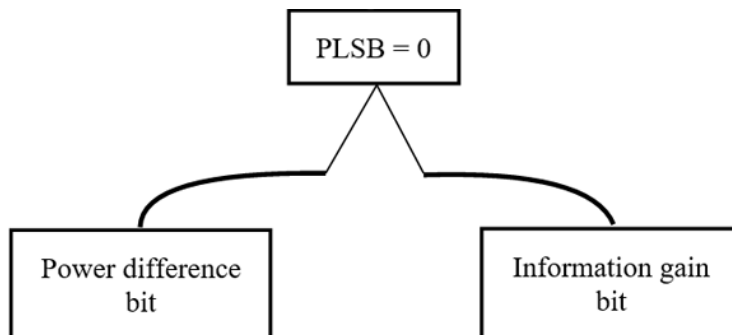
Further, data density is defined as a product of data buffer amount and information gain. How much information has been gained from the last round or it has turned negative and conserving data from that node now proves to be irrelevant. This is the way through which we restrict the selection of the number of cluster heads around for the mobile agent to retrieve the data to the sink.

The formulations are as follows:

$$P = D / \left(d \times t \right) \tag{1}$$

PEDAS measure is directly proportional to data density and power difference while inversely proportional to the time window for transmission as defined above. We can further breakdown the data density as:

A = amount of data processing
I = Information gain value from the previous round can be negative or positive and even zero

Putting in Eq. (1), we get:

$$P = \left(\frac{A \times I}{d \times t} \right)$$

Here, value (A/t) is termed as T is node-throughput. T is the particular node throughput:

$$T = \frac{A}{t}$$

The sink node having the capability of efficient processing to calculate the PEDAS value for different cluster heads and select those from where data needs to be fetched by the mobile agent-based on PEDAS values. Greater PEDAS values are prioritized for data gathering. Having multiple moving agents can make it more costly however more throughput can be realized. This process is repeatedly applied to every data value capture by that particular sensor node.

3.3 Phase-3: Data Fetching Through Mobile Agent

In the final phase, the mobile agent will fetch the data back to the sink from already selected cluster heads based on the second phase of our approach. Then mobile agent (MA) has been assumed to have enough power to route through the given paths. While calculating the path, each power efficient cluster head is included in the minimum spanning tree algorithm. Mobile Agent moves on the decided coordinates as decided by the sink node. The sink node registers the static path for that particular iteration by suggesting the coordinates of the cluster heads. Mobile agent's moves across all the selected cluster heads however it doesn't take or gather the data from any node in the first pass rather it informs nodes with a message signal to get their buffers ready when it would be retreating.

$$w = a \left(\frac{1}{disij} \right) + \left(1 - a \right) \left(P_j \right) \tag{2}$$

where $0 < a \leq 1$.

Table 3. Network variable

Variables	Meaning
n	Total number of sensor nodes in the network
G_n	The set of sensor nodes in the network
G_{CH}	The set of cluster heads
I_j	Information gain amount of node j
P_j	PEDAS value of node j
CH	Cluster Head
S_{CH}	The set of selected cluster head for data collection by Mobile Agent (MA)
SG_{CH}	The set of sorted CH in decreasing order of PEDAS
dis_{ij}	Distance between node i and j

Sink calculates the path cost function between the cluster heads using equation-2 as it has all the information about other sensor nodes.

Algorithm 1: Calculation of PEDAS value and selecting cluster heads for data transmission

```
1:  ∀ node j in G_n do
2:       P_j←0
3:  ∀ node j in G_n do
4:       calculate P_j By Equation - 1
5:  sort CH in decreasing order of PEDAS
6:  SG_CH←CH in sorted order
7:  S_CH ← sink
10: ∀ node v_j in SG_CH do
11:      if I_j >0
12:             S_CH←S_CH+v_j
```

Algorithm 2: Path Planning

```
1: PATH_CH←sink
2: G←S_CH
3: while ∃ (u∈PATH_CH, v∈G) do
4:      find max w(u,v) By Equation - 2
5:      PATH_CH←PATH_CH+v
6:      G←G-v
7: end while
```

4. RESULT ANALYSIS

In this study, we simulate the entire sensor network on throughput, average energy utilization, alive node, and variance of energy. Due to simplicity, this simulation is performed on two-dimensional space. We are using a grid space of 200 x 200 m², within this grid sensor nodes are randomly deployed. To simulate this study and compare its results with previous work EEDAC and ATL algorithm, we are using 150 sensor nodes. At the beginning of the simulation, each node carries an equal amount

of energy as 2J. At any time of simulation, if the energy of any node becomes zero, it will not be considered in the result. The computing power of each sensor node is the same for the alliteration.

In the beginning, we implement our approach using NS3. EEDAC and ATL were implemented later using the same mess topology. We have analyzed the results on various parameters.

Figure 3 depicts the throughput between our approach and previous work EEDAC and ATL. Throughput between the source senor node and sink node is a vital parameter to check the stability of the entire network. The result proves that our approach provides more stability than EEDAC and ATL.

Figure 4 stablishes a mapping between PEDAS and previous research study EEDAC, ATL in varying simulation time. The variance of energy elaborates the distance from the average energy of the entire network. Our approach has lesser variance than EEDAC and ATL. It means energy is well balanced in the entire network.

Figure 5 establishes a relationship to count the alive node among algorithms. To validate the lifetime of the sensor network, we are using 150 randomly deployed sensor nodes. By results, we can say that in the proposed PEDAS algorithm data aggregation was well managed. In previous algorithms, almost 50% of the node are dead after the simulation.

The above graph compares average energy utilization in the varying simulation time. Clearly, compared to ATL scheduling, one of the modern approaches to network lifetime conservation, the PEDAS measure is showing approximately 31% less energy utilization to run the entire sensor network. PEDAS curve grows less steep as compared to EEDAC and

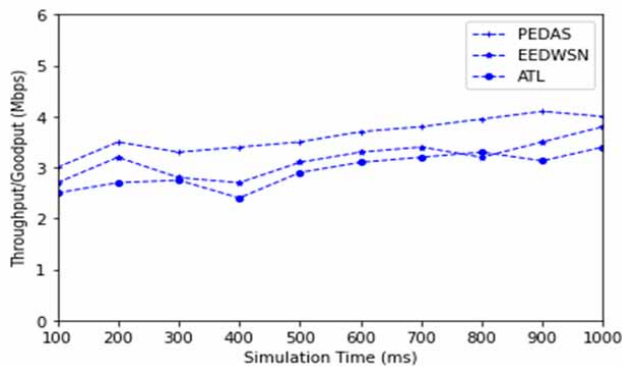**Figure 3. Throughput comparison among PEDAS, EEDAC, and ATL**

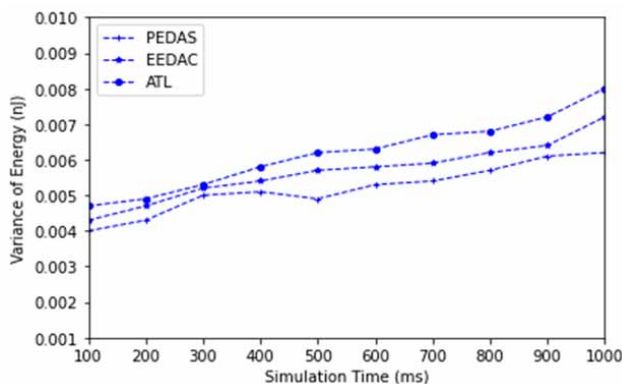**Figure 4. Variance of Energy in different simulation time**

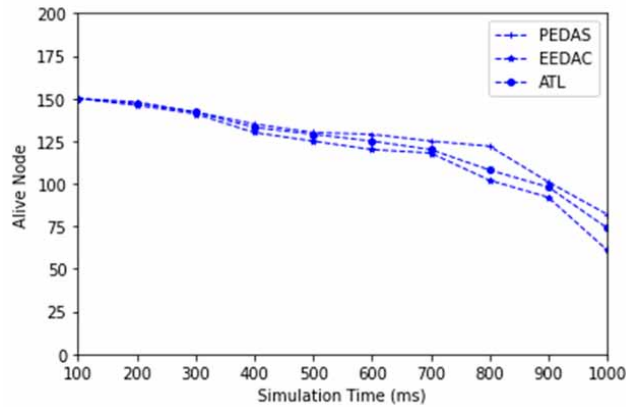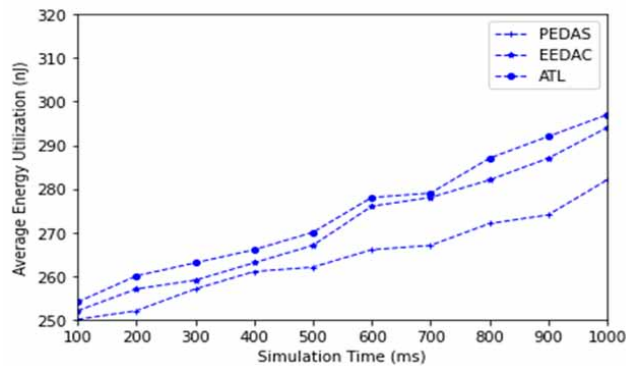Figure 5. Total number of an alive node by varying simulation time

Figure 6. Average energy utilization by the entire network

ATL scheduling with an increased simulation time. Therefore, we can conclude, PEDAS measures optimizing the network lifetime for the longer sessions. This research study efficiently used the data aggregation scheme. In summary, our algorithm has improved the lifetime of the entire network and provides a stable and robust network while comparing it with EEDAC-WSN and ATL schemes.

5. CONCLUSION

This study provided a technical overview of data aggregation in the Wireless Sensor Network. This study focused on information acquisition, efficient clustering, and picking the heads based on the PEDAS value. PEDAS measure is directly proportional to data density and power difference while inversely proportional to the time window for transmission According to the findings, all approaches sought to increase network throughput and reduce energy consumption. The proposed procedures are aimed at lowering the system's energy consumption. They, on the other hand, disregard data security. The mobile agent-based procedures improve data transmission delay and energy consumption while

also addressing device attributes. Furthermore, a review of existing methodologies reveals that, despite a variety of research approaches in the WSN data aggregation sector, security, reliability, heterogeneity, and data transmission delay remain unsolved problems in this field. As a result, future study should concentrate on all of these issues at the same time.

FUNDING AGENCY

REFERENCES

Al-Humidi, N., & Chowdhary, V. (2019, May). Lightweight Data Transmission Scheme Based on Data Aggregation Technique in Wireless Sensor Networks. *Proceedings of International Conference on Communication and Information Processing (ICCIP)*.

Al-Karaki, J. N., Ul-Mustafa, R., & Kamal, A. E. (2004, April). Data aggregation in wireless sensor networks-exact and approximate algorithms. In *2004 Workshop on High Performance Switching and Routing, 2004. HPSR.* (pp. 241-245). IEEE. doi:10.1109/HPSR.2004.1303478

Arena, F., & Pau, G. (2020). An overview of big data analysis. *Bulletin of Electrical Engineering and Informatics*, *9*(4), 1646–1653. doi:10.11591/eei.v9i4.2359

Asemani, M., & Esnaashari, M. (2015). Learning automata based energy efficient data aggregation in wireless sensor networks. *Wireless Networks*, *21*(6), 2035–2053. doi:10.1007/s11276-015-0894-3

Bongale, A. M., Nirmala, C. R., & Bongale, A. M. (2020). Energy efficient intra-cluster data aggregation technique for wireless sensor network. *International Journal of Information Technology*, 1-9.

El Fissaoui, M., Beni-Hssane, A., & Saadi, M. (2019). Energy efficient and fault tolerant distributed algorithm for data aggregation in wireless sensor networks. *Journal of Ambient Intelligence and Humanized Computing*, *10*(2), 569–578. doi:10.1007/s12652-018-0704-8

Giliberto, M., Arena, F., & Pau, G. (2019). A fuzzy-based Solution for Optimized Management of Energy Consumption in e-bikes. *Journal of Wireless Mobile Networks, Ubiquitous Computing and Dependable Applications*, *10*(3), 45–64.

Gupta, S., & Khan, G. (2016, November). MHCDA: A proposal for data collection in Wireless Sensor Network. In *2016 International Conference System Modeling & Advancement in Research Trends (SMART)* (pp. 193-196). IEEE. doi:10.1109/SYSMART.2016.7894517

Krishna, M. B., & Vashishta, N. (2013, September). Energy efficient data aggregation techniques in wireless sensor networks. In *2013 5th International Conference and Computational Intelligence and Communication Networks* (pp. 160-165). IEEE. doi:10.1109/CICN.2013.143

Latha, A., Prasanna, S., Hemalatha, S., & Sivakumar, B. (2019). A harmonized trust assisted energy efficient data aggregation scheme for distributed sensor networks. *Cognitive Systems Research*, *56*, 14–22. doi:10.1016/j.cogsys.2018.11.006

Lohani, D., & Varma, S. (2016). Energy efficient data aggregation in mobile agent based wireless sensor network. *Wireless Personal Communications*, *89*(4), 1165–1176. doi:10.1007/s11277-016-3310-0

Manoharan, V., & Chinnadurai, M. (2016). Energy Efficient Wireless Sensor Network Lifetime Maintenance using LPEDAP. *Energy*, *5*(01), 7–11.

N, K. K., & A, B. (2020). Real time energy efficient data aggregation and scheduling scheme for WSN using ATL. *Computer Communications*, *151*, 202–207. doi:10.1016/j.comcom.2019.12.027

Padmaja, P., & Marutheswar, G. V. (2018). Energy efficient data aggregation in wireless sensor networks. *Materials Today: Proceedings*, *5*(1), 388–396. doi:10.1016/j.matpr.2017.11.096

Ramachandran, N., & Perumal, V. (2018). A Honey-Hive based Efficient Data Aggregation in Wireless Sensor Networks. *Journal of Electrical Engineering & Technology*, *13*(2), 998–1007.

Roy, N. R., & Chandra, P. (2019, April). EEDAC-WSN: Energy Efficient Data Aggregation in Clustered WSN. In *2019 International Conference on Automation, Computational and Technology Management (ICACTM)* (pp. 586-592). IEEE. doi:10.1109/ICACTM.2019.8776679

Sasirekha, S., & Swamynathan, S. (2017). Cluster-chain mobile agent routing algorithm for efficient data aggregation in wireless sensor network. *Journal of Communications and Networks (Seoul)*, *19*(4), 392–401. doi:10.1109/JCN.2017.000063

Shanmukhi, M., & Ramanaiah, O. B. V. (2015, February). *Cluster-based comb-needle model for energy-efficient data aggregation in wireless sensor networks. In 2015 Applications and Innovations in Mobile Computing (AIMoC)*. IEEE.

Shobana, M., Sabitha, R., & Karthik, S. (2021). Cluster-Based Systematic Data Aggregation Model (CSDAM) for Real-Time Data Processing in Large-Scale WSN. *Wireless Personal Communications*, *117*(4), 2865–2883. doi:10.1007/s11277-020-07054-2

Sinha, A., & Lobiyal, D. K. (2013). A multi-level strategy for energy efficient data aggregation in wireless sensor networks. *Wireless Personal Communications*, *72*(2), 1513–1531. doi:10.1007/s11277-013-1093-0

Yestemirova, G., & Saginbekov, S. (2018, May). Efficient data aggregation in wireless sensor networks with multiple sinks. In *2018 IEEE 32nd International Conference on Advanced Information Networking and Applications (AINA)* (pp. 115-119). IEEE. doi:10.1109/AINA.2018.00029

Yuvaraj, R., Chandrasekar, A., & Jothi, S. (2019). Time orient LEAD based polling point selection algorithm for efficient data aggregation in wireless sensor networks. *Cluster Computing*, *22*(2), 3339–3346. doi:10.1007/s10586-018-2166-3

Zhang, J., Lin, Z., Tsai, P. W., & Xu, L. (2020). Entropy-driven data aggregation method for energy-efficient wireless sensor networks. *Information Fusion*, *56*, 103–113. doi:10.1016/j.inffus.2019.10.008

Abhishek Bajpai received his master's in Wireless Communication and Computing from Indian Institute of Information Technology, India and the Bachelor of Technology (First Class) in Information Technology from India. He is currently an Assistant Professor of Computer Networks at the Department of Computer Science and Engineering, Government Engineering College, Kannauj, UP, India. Prior to joining GEC, he was a Research Student at IIIT-A in the Wireless Sensor Network Laboratory. His current research interests include Wireless Computing, MachineLearning and Cloud Computing. Mr Bajpai has published many technical papers in major international journals and conferences of repute. He serves as a Review Member of several journals, including Ad-hoc Networks Elsevier, Journal of supercomputing, Springer, OJCST, IJAIEM. He is in member of the Editorial Board of JACOTECH, International Journal of Distributed Computing and Technology, Journals Pub. He is an active member of ACM, Computer Society of India, life member of the Indian Science Congress, CSTA-USA. He is serving as a technical program committee member of various International Conferences.

Mansi Chaurasia is currently pursuing Bachelor in computer science and engineering from Abdul kalam technical university. Her current research interests include data aggregation in wireless sensor networks.