EEDAC-WSN: Energy Efficient Data Aggregation in Clustered WSN

Nihar Ranjan Roy
Assistant Professor
G. D. Goenka University
Haryana, India
niharranjanroy@gmail.com

Pravin Chandra
Professor
Guru Gobind Singh Indraprastha University
New Delhi, India
chandra.pravin@gmail.com

***Abstract*—One of the primary challenges in Wireless Sensor Networks (WSNs) is energy efficiency. In the last decade WSNs have found its application in wide range of applications producing huge amount of data. Energy efficiency of a clustered WSNs can be improved, if size of data being transmitted in the network is reduced. In this paper, we present an energy efficient data aggregation scheme for clustered wireless sensor network (EEDAC-WSN). It reduces intra-cluster communications by allowing cluster member nodes to send small sized control frames followed by relatively detailed frames from nodes selected by the cluster head node. The proposed work can dovetailed with any clustering scheme, we have used it with LEACH for the purpose of simulation. The results obtained are substantial in terms of network stability period and lifetime.**

Keywords—Data Aggregation, Clustering, Wireless Sensor Network

I. Introduction

Wireless Sensor Networks (WSNs) consists of large number of re- sources constrained, inexpensive nodes capable of sensing the physical environment of deployment, cooperate with peer nodes and send the data to base station node. WSNs have found their application in broad range of sensing applications such as: habitat monitoring, healthcare, smart cities, border surveillance, and structural monitoring.

As WSN nodes are severely resource constrained nodes. They have small size battery to power all its internal units, limited processing power and low power communication unit. These constrains along with rise in new data intensive applications necessitate need for novel energy conservation strategies. Out of several solutions pro- posed by researchers in the last two decades, clustering is one of the prominent options. Clustering offers certain advantages such as scalability, robustness and energy efficiency. The need and benefits of data aggregation has been pointed and confirmed by early research as in [1] – [3]. Energy efficient data aggregation is a natural choice of an energy efficient WSN, but it is still relatively less exploited area. In this paper, we have presented an energy efficient data aggregation scheme for clustered WSN. Work presented here, reduces the number of bits transmitted during intra-cluster communications. Nodes here send a relatively small size meta-data frame followed by data frame from selected node.

The remainder of this paper is organized as follows. Section II discusses the related work. section III describes network layout. Section IV explains the proposed work. Section V discusses simulation and results, and section VI finally concludes the work.

II. Related Work

Many clustering protocols promising energy efficiency have been proposed by researchers in the past. Some of the popular techniques are LEACH [4], HEED [5], PEGASIS [6], and WCA [7]. These techniques select cluster heads on the basis of criteria like randomized approach in LEACH, residual energy in HEED, and node degree in WCA. Many descendants of these protocols have also been developed in the recent years.

Many data aggregation schemes targeting WSN have also been proposed in the past. Sarangi et al. [8] have presented work on data aggregation using Ant Colony Optimization, ShivKumar et al. [9] work is inspired from GSTEB [10] where root of the tree performs as cluster head. Mottaghi et al. [11] have used mobile sink and rendezvous nodes for data aggregation. Yuan et al. [12] have used data density and spatial correlation between the data nodes and its neighbours. There are different techniques of data aggregation as per the survey work done by [13], [14]. In-network data aggregation is popular data aggregation technique a survey of such work is available in [15]. A general analysis, survey and classification of latest data aggregation techniques in WSN is available in [16].

Chen et al. [17] proposed a meta-data-based data aggregation scheme for clustered sensor networks which sends meta data packets to the Cluster head node followed by data from by data source node. The disadvantage of this approach is that it collects data from only one data source so it cannot guarantee reliability. Secondly it cannot detect multiple events in a cluster with small event radius model. Thirdly, cluster head node, cannot send a summarized data about its cluster. It also fails to clearly explain the selection process of data source node for data transmission on the basis of data received in meta-data frame. These disadvantages are overcome in our work presented here.

III. Network Layout

Let us consider a WSN of N sensor nodes, uniformly and randomly deployed in a square region of M X M. The cost of communication from node $N_i \rightarrow N_j$ is same as communication from $N_j \rightarrow N_i$. All the sensed data must reach the base station (BS) (also known as sink) in some or the other form.BS node is not resource constrained and can be placed anywhere in the network, on the boundary of the network or outside the network of interest. A cluster in WSN can be viewed as a circular region with radius ‘r’. Inside this cluster there are some sensor nodes, among which a cluster head (CH) node is chosen. CH node is assumed to be at the center of the cluster region or in other words cluster region is formed around the CH. All other nodes except CH node inside the cluster are hereby referred as non- cluster head nodes (NCH).

CH node coordinates with the member nodes of the cluster, collects sensed data from them and does some elementary processing before forwarding it to the sink. Thus, it can be seen that CH node acts as a representative on behalf of its cluster. Normally sink nodes are located at larger distance in comparison to distance between CH and NCH nodes of a cluster, which means CH nodes are responsible for long distance communication.

A. *Radio Communication Model*

There are several radio communication models proposed by researchers for WSN. Here, we have used one of the most popular models proposed by Heinzelman et al. [4]. As per this model, the energy required to run the transmitter and receiver circuitry are ETx−amp and ERx−elec respectively. Energy required to transmit ‘l’ bits of data over a distance of ‘d’ meters is

$$E_{Tx}(l,d) = E_{Tx}(l) + E_{Tx-amp}(l,d) \quad (1)$$

If the transmission distance is small, then free space propagation model will be used as below

$$E_{Tx-amp}(l,d) = lE_{elec} + l\epsilon_{fs}d^2 \quad if\ d < d_0 \quad (2)$$

else two-ray ground model will be used as below

$$E_{Tx-amp}(l,d) = lE_{elec} + l\epsilon_{mp}d^4 \quad if\ d \geq d_0 \quad (3)$$

where d_0 is known as crossover distance.

Similarly, energy dissipated by receiver is

$$E_{Rx}(l) = E_{Rx-elec}(l) = lE\_(elec) \quad (4)$$

B. *Event Source Model*

Process of data aggregation can be affected by number of events, position of events and communication path. In an event driven WSN there are primarily three types of event models: Even radius model, Random source model [18] and Hybrid Model [17].

1) Event-Radius Model: In this approach an event can be considered as a point source, where all the nodes within a radius of S are called source nodes as they can sense this event.
2) Random Source Model: In this model, randomly some nodes are selected as source of the event.
3) Hybrid Model: In this model [17], a mixed approach is followed. First a random source model is used to generate event source nodes, secondly few of these nodes are selected as data sources.

In this paper, we follow an event radius model, where all the nodes within a radius of S from the event sense it. We have referred these nodes as data source nodes.

IV. Energy Efficient Data Aggregation in Clustered WSN - Proposed Work

We follow a cluster-based approach, where after initialization sensor nodes communicate with each other and form clusters. The clustering process is broken into two phases: setup-phase and steady state phase. It is assumed that setup phase is much smaller then steady phase. Data transmission takes place during steady phase. During this phase nodes follow a TDMA based scheme to avoid collision and maintain a Sleep-Wake cycle.

A. *Scheme Operation*

In a traditional data aggregation scheme adopted by LEACH and its descendants protocols, all the member nodes send the sensed data to the CH node as per a TDMA schedule in a periodic fashion. In our case, NCH nodes send a control signal (meta-data) packet to the CH node rather than sensed data. Once all the meta-data packets are received, CH decides from whom to gather the complete event data as shown in the fig. (1). Before we explain data collection phase in detail, below are some of our assumptions:

1) All the nodes have equal capabilities.
2) All the nodes are capable of communicating to the BS, if needed.
3) Nodes store average of data transmitted since beginning or after reset of the network.
4) During initialization all the nodes have value of Δd, a small change in data, which should be reported.
5) All the nodes are time synchronized.

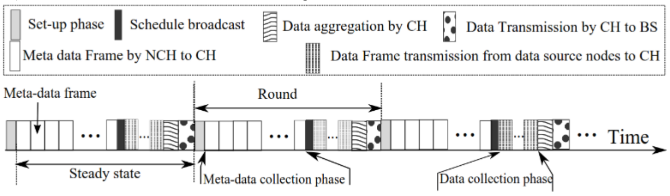

Fig. 1. TDMA Scheme for Data Aggregation

Let us assume that, the average of previous data sensed by node N_i till round r is $D_{avg,i,r}$, $E_{i,r}$ is its residual energy and current sensed data is $d_{i,r}$ having time stamp $t_{i,r}$. In a meta data frame a node N_i sends its node id, $E_{i,r}$,$t_{i,r}$ and an integer value $V_{i,r}$ indicating, it has new value to send or not. Where $V_{i,r}$ is estimated using following rule

$$V_{i,r} = 0 \ \text{ if } \ |d_{i,r} - D_{avg,i,r}| < \Delta d \quad (5)$$

else an integer value greater than 0 will be sent, assuming higher the value more significant is the data.

$$V_{ir} = \left\lceil \frac{|d_{ir} - D\_(avg,i,r)|}{\Delta d} \right\rceil \quad if\ \left|d_{ir} - D_{avg,i,r}\right| > \Delta d \quad (6)$$

As shown in fig. (1), a CH node receives meta-data frames from its member nodes which contains mandatory fields like (node id, $E_{i,r}$, $t_{i,r}$, $V_{i,r}$). Once all the meta-data frames are received, CH may decide to receive data from those nodes who have sent $V_{i,r}$ > 0. Before data collection, CH node broadcasts a time schedule for its NCH nodes (data source nodes). Data source nodes listen to this broadcast message and transmit data at the specified slot. Here, it is assumed that length of meta-data frame l_{mdf} is much smaller than length of data frame l_{df}. Once all data is received, CH node may perform data aggregation and send data to BS or towards BS node as shown in fig. (2).

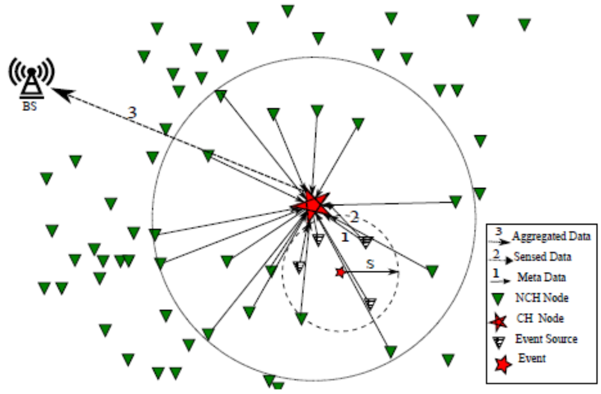

Fig. 2. Data aggregation model, NCH nodes and Data Source nodes send meta-data frame

During data collection phase the CH node may either choose to collect data from all the data source nodes or data source node with the max or min value. Data source node with highest $V_{i,r}$ carries the highest change in data. If CH decides to receive data from all the data source nodes, then it would be able to send a summarized data from the cluster to the BS. In case the CH node wants to collect data from the node reporting maximum change and there are multiple nodes with same $V_{i,r}$ value, then

1) If the CH wants to collect data from the data source node which first sensed the value, it may do so by choosing a data source node on the basis of time stamp $t_{i,r}$.
2) If the CH wants to collect data from the data source node which has higher residual energy, it may do so by choosing a data source node on the basis of $E_{i,r}$.

In another approach (here by called as EEDAC-WSN-Silent), NCH nodes whose $V_{i,r} = 0$, may opt not to send even meta-data frame, as they do not have substantial data change to report and may not listen to the next broadcast schedule for data transmission. A disadvantage of this approach could be that the CH node may not know how many of its member nodes are alive after cluster formation phase. This approach further reduces intra-cluster communication cost.

A recent work by [19], suggests that in a clustered WSN, it is not necessary that all the nodes are part of a cluster, under certain scenarios they may directly communicate with the base station. As shown in fig. (3) for a sensor field with dimensions 100 X 100 and base station located at (50,150), it can be seen that nodes falling in the grey circle are supposed to communicate directly to the BS. In such a case we suggest that only data source nodes should send data to the BS and BS performs the responsibilities of a CH. For such nodes, after every CH selection process, BS will remain their cluster head.

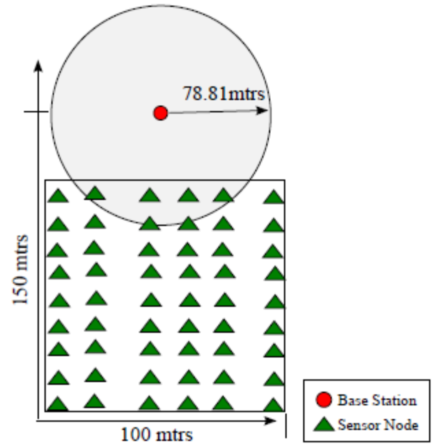

Fig. 3. Scenario where nodes are not part of a cluster [19]

In this section we analyze the energy consumption for the proposed model. Assuming that N sensor nodes are uniformly and randomly deployed in a sensing filed with dimension M X M. All the nodes are homogeneous and have initial energy E_0 and BS has infinite energy. At an average, we have k number of clusters in the network. Intra cluster communications follow free space propagation and inter cluster communication follow multi-path fading.

Number of data source nodes N_{ds} in the network can be estimated as

$$N_{ds} = \pi S^2 \frac{N}{M^2} \qquad (7)$$

Inside a cluster there are $\left(\frac{N}{k} - 1\right)$ member nodes excluding the CH node. These member nodes consist of data source nodes and non-data source nodes. Energy consumption by a CH node inside a cluster where an even has occurred/ a data source node exists is estimated using the following expression

$$E_{CH} = \left(\frac{N}{k} - 1\right) E_{Rx}\left(l_{mdf}\right) + E_{Tx}(l_s, d_{CH-sn}) + N_{(ds)}E_{DA}l_{df} + N_{(ds)}E_{Rx}\left(l_{df}\right) + E\_(Tx)\left(l_{df}, d_{BS}\right) \qquad (8)$$

which includes energy dissipated in:

1) receiving l_{mdf} bits of meta-data message from NCH nodes of the cluster.
2) broadcasting the schedule of data collection to data source members with message length ls bits. Here d_{CH-sn} denotes distance of the farthest member node from the CH node.
3) receiving l_{df} bits of data frame message from N_{ds}
4) aggregating data received from N_{ds} data source nodes, E_{DA} is the energy dissipated in aggregating per bit of data.
5) transmitting aggregated data to BS, located at a distance d_{BS} from the CH node.

If silent mode approach is followed, where non-data source nodes decide not to send the meta-data frame, above equation can be written as:

$$E_{CH} = N_{ds}E_{Rx}\left(l_{mdf}\right) + E_{Tx}\left(l_s, d_{CH-sn}\right) + N_{(ds)}E_{DA}l_{df} + E\_(Tx)\left(l_{df}, d_{BS}\right) \quad (9)$$

Energy dissipated by a data source node, is estimated using following expression

$$E_{DS} = E_{Tx}\left(l_{mdf}, d_{CH}\right) + E_{(Rx)l_s} + E_{Tx}\left(l_{df}, d_{CH}\right) \quad (10)$$

Similarly, energy dissipated by a non-data source node is estimated using

$$E_{NCH} = E_{Tx}\left(l_{mdf}, d_{CH}\right) \quad (11)$$

Where d_{CH} is distance from the node to CH node.

Optimum number of cluster is estimated using the following expression by [19].

$$K = \begin{cases} N \; : d_{BS} \leq t - distance \\ \sqrt{\frac{N}{2\pi}}\sqrt{\frac{\epsilon_{fs}}{\epsilon_{mp}}}\frac{M}{\sqrt{d_{BS}^4 - \frac{E_{elec}}{\epsilon_{mp}}}} \; : otherwise \end{cases} \quad (12)$$

where t-distance is the threshold distance.

$$t - distance = \left[\frac{1}{\epsilon_{mp}}\left(\frac{M^2\epsilon_{fs}}{2\pi N} + E_{elec}\right)\right]^{\frac{1}{4}} \quad (13)$$

V. Simulation Results and Discussion

The proposed Energy Efficient Data Aggregation in clustered WSN (EEDAC-WSN) can be dovetailed with any clustering protocol. In this paper, we have implemented it using LEACH [4]. One of the reasons for choosing LEACH is its popularity and simplicity. We have also implemented it using Roy & Chandra recent work on LEACH [19], because of significant contribution that under certain scenarios even in a clustered WSN, clustering may not be required. We have used 4 performance evaluation metrics FND (First node dead), LND (Last node dead), Number of packets transmitted to CH and BS. First node dead period from beginning is also known as stability period.

Under current experimental set-up, we have considered a sensing field of size 100 X 100 in which 100 sensor nodes are deployed uniformly and randomly. Base station is located at location (50,150) i.e., outside the sensing field.The initial size of the meta-data frame and broadcast frame is 1000 bits, and data frame is 4000 bits. Radio parameters, E_{DA}, d_0, k and initial energy of the nodes is same as that in [4].Event radius is 10 meters and number of events is 4. Lifetime comparison of LEACH, LEACH by Roy et al. [19], EEDAC-WSN and EEDAC-WSN-Silent is shown in fig. (4). From fig. (5) it is observed that the lifetime of our proposed work (EEDAC- WSN) is better than LEACH and Roy et al. work. It is further observed EEDAC-WSN-Silent is even better than the other three. With respect to original LEACH, EEDAC-WSN and EEDAC-WSN-Silent have increased stability by 17.67% and 23%.Similarly increase lifetime of EEDAC-WSN and EEDAC-WSN-Silent is by 160.39% and 171.89% respectively. Similarly a substantial improvement in throughput is observed as shown in fig.(6 and 7) and tabulated in table.(I). The values shown in all the tables hereafter are average of 30 simulations. The substantial improvement in packets being transmitted to BS is due to the fact that all the data source nodes whose distance to BS is less than the threshold distance communicate directly to the BS.

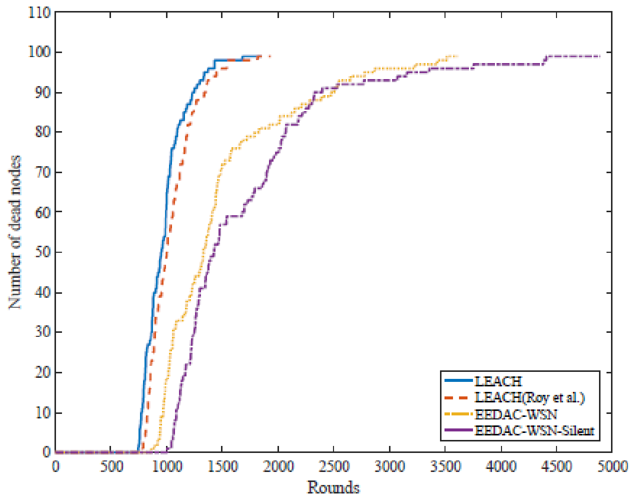

Fig. 4. Lifetime comparison

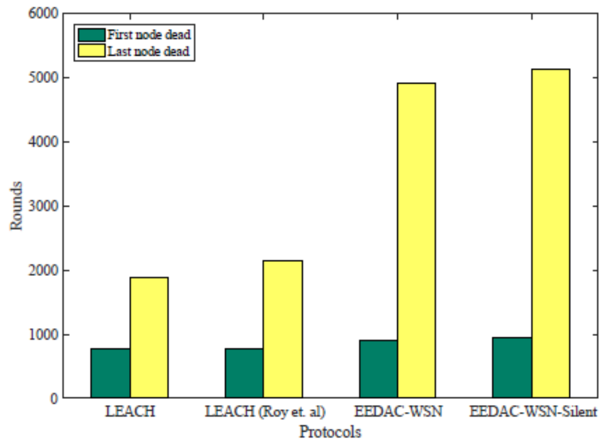

Fig. 5. Network stability and lifetime

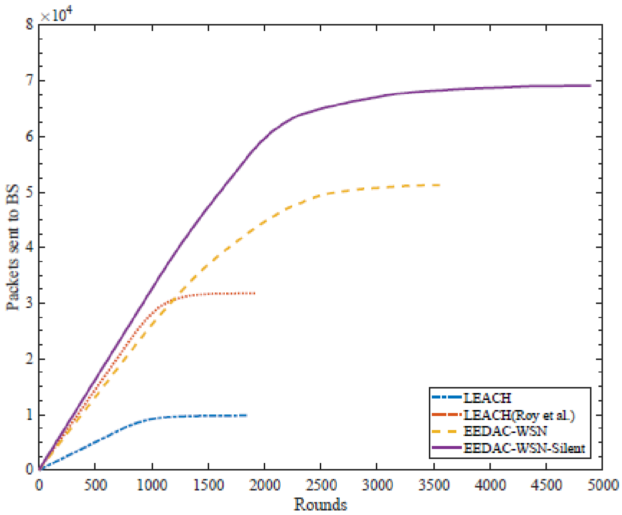

Fig. 6. Packets transmitted to base station

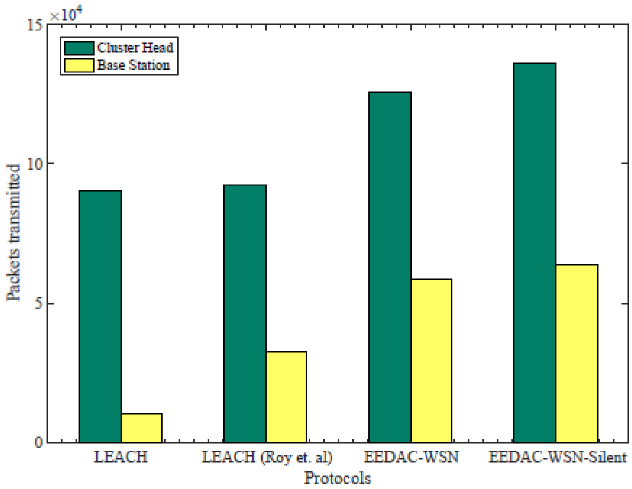

Fig. 7. Throughput: Number of packets transmitted to CH and BS

TABLE I. LIFETIME AND THROUGHPUT COMPARISON

	FND	LND	PktToCH	PktToBS
LEACH	772.3	1888.2	90396.3	10247.4
LEACH Roy et al.	779.4	2141.4	92657.7	32496.9
EEDAC-WSN	908.8	4916.7	125751.9	58816.5
EEDAC-WSN-Silent	950	171.88	136059.3	63560.4

We also studied the effect of change in the event radius and tabulated the results in table. (II). Effect of change in event radius on network stability and network lifetime is shown in fig. (8 and 9) respectively and tabulated in table. (II). It is observed that as event radius (S) increases more sensor nodes sense event(s), which leads to sending of more data frames to CH(s). As number of data frames being sent increases there is decrease in stability and lifetime due to faster rate of energy dissipation.

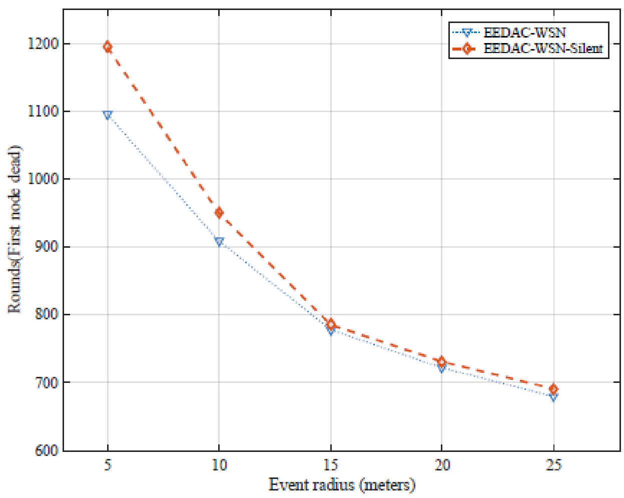

Fig. 8. Effect of event radius on stability period of network

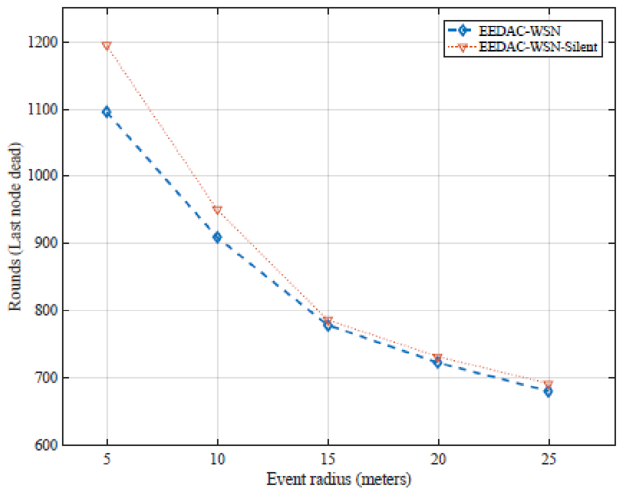

Fig. 9. Effect of event radius on network lifetime

TABLE II. EFFECT OF EVENT RADIUS (S) ON STABILITY PERIOD AND NETWORK LIFETIME

	FND					LND				
Event Radius	5	10	15	20	25	5	10	15	20	25
EEDAC-WSN	1095	909	778	722	680	6345	4917	3727	2980	2532
EEDAC-WSN-Silent	1195	950	786	731	691	6158	5134	3747	2973	2497

In order to study the effect of number of events in the sensing field. We simulated, for different values of number of events and tabulated the finding in table. (III). From fig. (10 and 11) it is seen that there is decrease in the network stability and lifetime due to increase in the number of events. The reason for this is, when number of events increase, more sensor nodes become data source nodes and transmission of data frames increases.

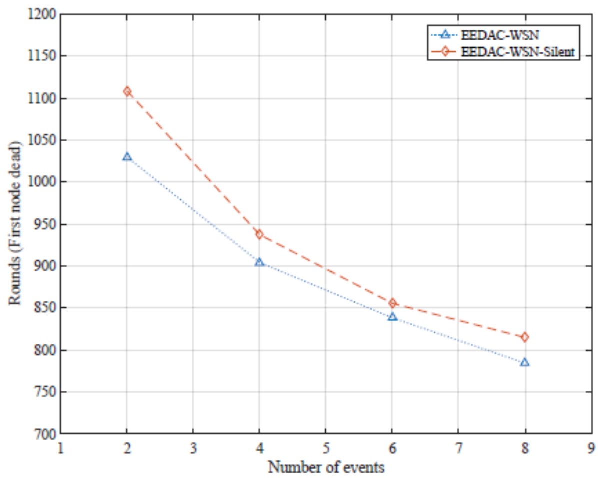

Fig. 10. Effect of number of events on stability of network

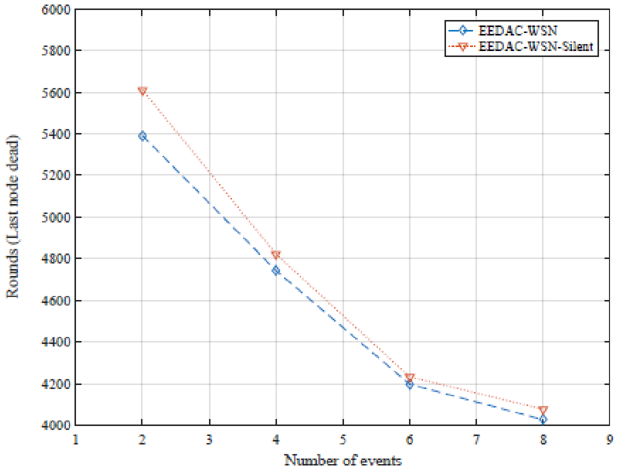

Fig. 11. Effect of number of events on network lifetime

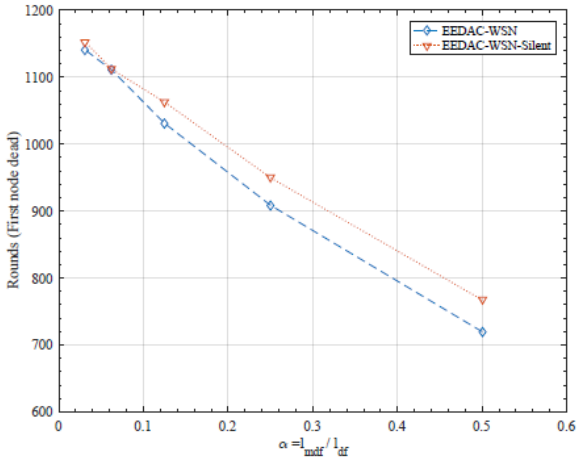

Fig. 12. Effect of on stability of network

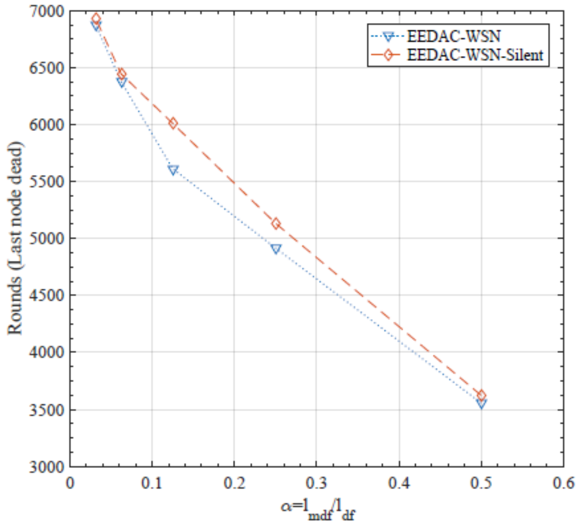

Fig. 13. Effect of on network lifetime

TABLE III. Effect of Number of Events on Network Stability and Lifetime

	FND				LND			
Number of events	2	4	6	8	2	4	6	8
EEDAC-WSN	1029.6	904.2	838.4	784.4	5392	4743	4196.6	4026.1
EEDAC-WSN-Silent	1108.5	937.4	855.6	815	5612	4822	4232.1	4076.3

Finally, we studied the effect of change in the ratio of length of meta data frames and data frames, defined as $\propto = \frac{l_{mdf}}{l_{df}}$. We took different values of l_{mdf} and recorded the observations in table. (IV) and plotted in fig. (12 and 13). It is observed that less is the value of α more is the stability period and network lifetime. This is because less number of bits will be transmitted during intra-cluster communications and energy will be saved.

TABLE IV. Effect of a on Network Stability and Lifetime

	FND					LND				
α	1/2	1/4	1/8	1/16	1/32	1/2	1/4	1/8	1/16	1/32
EEDAC-WSN	719.9	908.8	1031	1110.8	1140.8	3556	4917	5612	6375.5	6874.7
EEDAC-WSN-Silent	767.3	950	1063	1112	1152.6	3624	5134	6011	6442.7	6930.8

VI. Conclusion

In this paper, we have proposed an Energy Efficient Data Aggregation scheme for Clustered WSNs, EEDAC-WSN and its variant EEDAC-WSN-Silent. In this work we tried to reduce intra-cluster communication by sending smaller frames representing nodes status to send new data (min Δd change in comparison to average data sent before current round) followed by relatively detailed frame from node(s) as selected by CH. Simulations show substantial increase in stability period of the network and lifetime. We also studied the effect of influential factors: event radius, number of events and ratio of meta-data and data frame and found that the technique is consistent. The proposed approach can be dovetailed with any clustering technique, we have done here with LEACH and latest variant of LEACH by Roy et al.

References

[1] B. Krishnamachari, D. Estrin, and S. Wicker, "The impact of data aggregation in wireless sensor networks," in *null*, p. 575, IEEE, 2002.

[2] H. O¨. Tan and me. Ko¨rpeog˘lu, "Power efficient data gathering and aggregation in wireless sensor networks," *ACM Sigmod Record*, vol. 32, no. 4, pp. 66–71, 2003.

[3] K. Kalpakis, K. Dasgupta, and P. Namjoshi, "Maximum lifetime data gathering and aggregation in wireless sensor networks," in *Networks*, pp. 685–696, World Scientific, 2002.

[4] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "An application- specific protocol architecture for wireless microsensor

networks," *Wire- less Communications, IEEE Transactions on*, vol. 1, pp. 660–670, Oct 2002.

[5] O. Younis and S. Fahmy, "Heed: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," *IEEE Transactions on mobile computing*, no. 4, pp. 366–379, 2004.

[6] S. Lindsey and C. S. Raghavendra, "Pegasis: Power-efficient gathering in sensor information systems," in *Proceedings, IEEE Aerospace Con- ference*, vol. 3, pp. 3–3, IEEE, 2002.

[7] M. Chatterjee, S. K. Das, and D. Turgut, "Wca: A weighted clustering algorithm for mobile ad hoc networks," Cluster computing, vol. 5, no. 2, pp. 193–204, 2002.

[8] K. Sarangi and I. Bhattacharya, "A study on data aggregation techniques in wireless sensor network in static and dynamic scenarios," Innovations in Systems and Software Engineering, pp. 1–14, 2019.

[9] S. Shivkumar, A. Kavitha, J. Swaminathan, and R. Navaneethakrishnan, "General self-organizing tree-based energy balance routing protocol with clustering for wireless sensor network," Asian Journal of Information Technology, vol. 15, no. 24, pp. 5067–5074, 2016.

[10] Z. Han, J. Wu, J. Zhang, L. Liu, and K. Tian, "A general self-organized tree-based energy-balance routing protocol for wireless sensor network," IEEE Transactions on Nuclear Science, vol. 61, no. 2, pp. 732–740, 2014.

[11] S. Mottaghi and M. R. Zahabi, "Optimizing leach clustering algorithm with mobile sink and rendezvous nodes," AEU-International Journal of Electronics and Communications, vol. 69, no. 2, pp. 507–514, 2015.

[12] F. Yuan, Y. Zhan, and Y. Wang, "Data density correlation degree clustering method for data aggregation in wsn," IEEE Sensors Journal, vol. 14, no. 4, pp. 1089–1098, 2014.

[13] R. Rajagopalan and P. K. Varshney, "Data aggregation techniques in sensor networks: A survey," *Electrical Engineering and Computer Science*, 2006.

[14] K. Maraiya, K. Kant, and N. Gupta, "Wireless sensor network: a review on data aggregation," International Journal of Scientific & Engineering Research, vol. 2, no. 4, pp. 1–6, 2011.

[15] E. Fasolo, M. Rossi, J. Widmer, and M. Zorzi, "In-network aggregation techniques for wireless sensor networks: a survey," IEEE Wireless Communications, vol. 14, no. 2, pp. 70–87, 2007.

[16] N. R. Roy and P. Chandra, "Analysis of data aggregation techniques in wsn," in International Conference on Innovative Computing and Communication, ICICC-2019, Springer, 2019.

[17] H. Chen, H. Mineno, and T. Mizuno, "A meta-data-based data aggrega- tion scheme in clustering wireless sensor networks," in 7th International Conference on Mobile Data Management (MDM'06), pp. 154–154, IEEE, 2006.

[18] B. Krishnamachari, D. Estrin, and S. Wicker, "Modelling data-centric routing in wireless sensor networks," in IEEE infocom, vol. 2, pp. 39–44, 2002.

[19] N. R. Roy and P. Chandra, "A note on optimum cluster estimation in leach protocol," IEEE Access, vol. 6, pp. 65690–65696, 2018.