

An Effective Clustering Approach with Data Aggregation Using Multiple Mobile Sinks for Heterogeneous WSN

A. Muthu Krishnan¹ · P. Ganesh Kumar²

Published online: 2 August 2015
© Springer Science+Business Media New York 2015

Abstract Wireless Sensor Networks (WSNs) mostly uses static sink to collect data from the sensor nodes randomly deployed in the sensor region. In the static sink based approach, the data packets are flooded across the network to reach the mobile base station in multi-hop communication. Due to this, the static sink is inefficient in energy utilization. Recently, mobile sink are used for data gathering, has less energy utilization which in turn increases the network lifetime. Thus, the sink mobility has difficulties in finding the routing path for the data packets. This paper proposes an effective clustering approach with data aggregation using multiple mobile sinks for heterogeneous WSN. The proposed algorithm achieves network lifetime increases with limited energy utilization.

Keywords Wireless Sensor Network (WSN) · Data aggregation · Clustering

1 Introduction

WSN is used in vast application domain such as military application, health-care, smart home and environmental monitoring [1–4]. In WSN, numerous small sensor nodes are randomly deployed in the sensing area. The sensor nodes are battery-powered, constrained with limited energy [5, 6]. Since, WSNs are widely used in application area such as target tracking, battlefield monitoring, environmental monitoring, it is impossible to replace the battery power of sensor nodes. Thus the power consumption of sensor node is the main constraint in WSN, which decides the network lifetime. The main functionalities of sensor node comprises of sensing, computation and wireless communication. In sensor nodes, the communication utilizes more power than sensing and computation.

✉ A. Muthu Krishnan
muthuathi@gmail.com

¹ Department of ECE, Anna University Regional Centre Madurai, Madurai, India

² K.L.N College of Engineering, Pottapalayam, Sivagangai, India

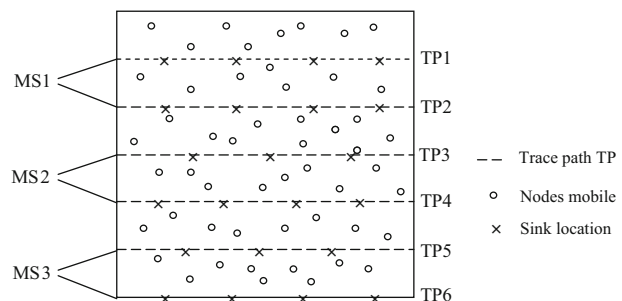
Thus, many routing protocols are introduced in WSN to reduce the amount of transmission for subsequent power reduction. In WSN, the data are flooded from the sensor node to the base station located at the center. If the sensor node is located far away from the base station, then the data packet has to traverse long distance which leads to more power consumption. Further, the sensor located around the base station referred as bottle neck zone experience heavy traffic flow and depletes the node's battery energy faster. Thus the nodes in bottle neck zone die faster which affects the network lifetime and network connectivity.

In order to avoid the power consumption due to long relaying of data packet, sink mobility based data gathering are introduced in WSN [7, 8]. The sink mobility significantly minimizes the number of hops the data packet traverses to reach the sink, which reduces the power consumption in data gathering. The sink mobility delays the transfer of data packet from any particular node to the sink, whereas the critical data requiring immediate action can be broadcasted by the sensor node to travel in multi-hop communication towards the sink. Due to the latency in data transfer, the sink mobility scheme is generally used in non-critical applications such as environmental monitoring [9–11]. The sink mobility achieves significant network lifetime improvement based on the relationship between mobility and data transfer delay [12–15].

2 Related Work

There are many recent work available on mobile sink based WSN. In directed diffusion [16], data centric routing based on attribute-based naming is used, where the mobile sink broadcast interest for data-queries. In Two-Tier Data Dissemination (TTDD) [17] uses grid based approach; the whole network is split to small grids. The sensor node close to the grid point act as dissemination node has the location information for routing. The mobile sink sends the data queries to the nearest grid and it is routed to other grid points, the node reply through reverse path. The TTDD based approach has less control overhead compared to other protocol. In [18], an efficient Data Driven Routing Protocol (DDRP) provides route learning by exploiting the wireless medium broadcast feature, where the data packet has an option called recording for each listener to learn the route to mobile sink. In [19], a Mobility-Assisted Data Collection model (MADC) provides a theoretical model for throughput capacity and lifetime based on parameters such as sink velocity, number of mobile sinks and travelling path. In [20], a Hierarchical Cluster-based Data Dissemination scheme (HCDD) is proposed where routing is performed

Fig. 1 Proposed architecture



without expensive GPS device for location information. Further several types of mobile sink are used to achieve the mobility for efficient coverage of sensor network. For Random movement of mobile sink, entities such as human, vehicles and animal are used. In [21], humans are used as mobile entity. In [22], the vehicles such as buses which move in fixed paths are used as mobile entity. In [23], whales are used as mobile entities used for data communication in WSN.

3 Proposed Work

The proposed method uses a novel data gathering based on clustering approach using multiple mobile sink for WSN shown in Fig. 1. The proposed method has several stages in data gathering they are (1) sink location trace path (2) Cluster formation stage (3) Data aggregation. If a single mobile sink is used in data gathering, the retrace of the same path is delayed for more time which increases the latency in data transfer. In the proposed method, multiple mobile sink is used in data gathering which retraces the same path with minimum delay. The sensor region is covered by several trace paths and the trace paths are across the region, a single mobile sink is used for pair of trace path.

The proposed method uses cluster based data gathering approach. The proposed method adopts Heterogeneous WSN, such that the high energy nodes are used as Cluster Head (CH). The CH receives the message from all the nodes in cluster, buffers the message and transmits the message to the mobile sink once it comes into contact. Since CHs experience heavy traffic flow, its energy gets depleted faster which results in loss of connectivity. In order to reduce the amount of traffic, Data Aggregation (DA) [24, 25] is used in the sensor nodes. If multiple nodes send data to the CH at the same time in the cluster, then data collide at the cluster. This can be overcome by allocating a TDMA slot to all nodes in the cluster by the CH. The heterogeneous node located along the trace path not selected as CH directly communicate with the mobile sink while it come into contact. This is achieved by proper moving speed of the mobile sink.

3.1 Sink Location Registration

Initially, the sink moves along the trace path and it halts at various location to initiate route discovery process. In each location, the sink broadcast the route discovery message called INIT message. The heterogeneous nodes which receive the INIT message, having energy level greater than the threshold level respond to mobile sink through a reply message having sender ID and its energy level. The mobile sink selects the node with high energy level as CH and sends the start message with the receive node ID to initiate the cluster formation. Further, the mobile sink evenly distributes the CH selection among the possible heterogeneous nodes at each location. This helps to avoid the energy depletion of particular heterogeneous node. When the mobile sink retraces the same path, the CH transmits the data of all sensor nodes to mobile sink.

The mobile sink contains the information of the previous CHs. Let W_i is the weight of the heterogeneous node S_i based on the past election as CH. Such that the heterogeneous node which is recently elected as CH has low weight and the heterogeneous node not recently elected as CH has more weight. Let ER_{CH} denotes the CH election record which contains the information of previous CH. E_i is the energy level of the heterogeneous node. The heterogeneous node having maximum weight is selected as Cluster Head.

$$W_i \propto (ER_{ch}(S_i) / \text{Max}(E_i)) \quad (1)$$

$$CH_{SN_Curr} = \max(W_i) \quad (2)$$

3.2 Cluster Formation Stage

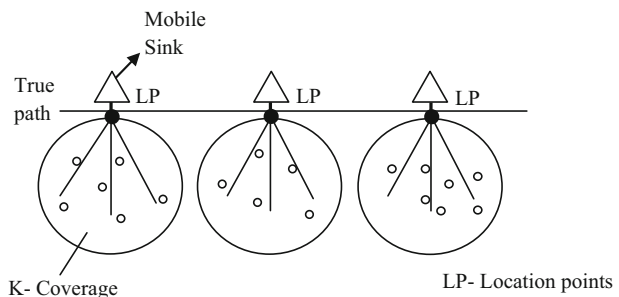
When the heterogeneous node receives the start message from the mobile sink, it becomes CH and starts the cluster formation process. The CH transmits its ID as the sender ID in small beacon message to the subsequent nodes. The Cluster is formed based on K-coverage scheme; due to the cluster contains k sensors to cover the region. Since the K-coverage scheme has better coverage, thus the proposed algorithm covers the sensor region efficiently. The sensor node which receives the message, access the channel potential using the message based on parameters such as Received Signal Strength Indicator (RSSI). Since the channel is symmetric, the received message with good RSSI indicates the channel is in good condition. Thus, the sensor node having high signal strength will join the cluster by sending the ACK message to the Cluster Head. The ACK message contains sender ID same as node ID along with the CH ID of the corresponding cluster, since the same sensor node will receive REQ messages from multiple CH. In order to avoid the data collision in CH, as data has been received from multiple sensor node. In the proposed method, the CH employs TDMA time slot to all sensor nodes in the cluster. Figure 2 shows the cluster architecture. Figure 3 shows the structure of data packet.

The CH uses equal time interval τ for all the sensor nodes. The CH informs the time schedule to the sensor node by transmitting a unique message, which contains the ID of sensor node in the order of time schedule, the schedule time interval and initial start time of TDMA time slot. The first sensor node in the order starts to transmit the data in the initial start time and completes within time interval. All other node begins the transmission in the time schedule calculated from the initial start time using the order as offset.

3.3 Data Aggregation

In the proposed method, the data from the sensor node are forwarded to the mobile sink only in its allocated time slot. Since the cluster contains more number of sensor nodes, it creates overhead in CH to transmit a large amount of data to the mobile sink. This overhead is avoided by employing data aggregation on the sensor nodes. The data aggregation also called as data fusion is application dependent. The sink mobility based approach is mostly used for non critical application such as environmental studies. In such

Fig. 2 Cluster Formation



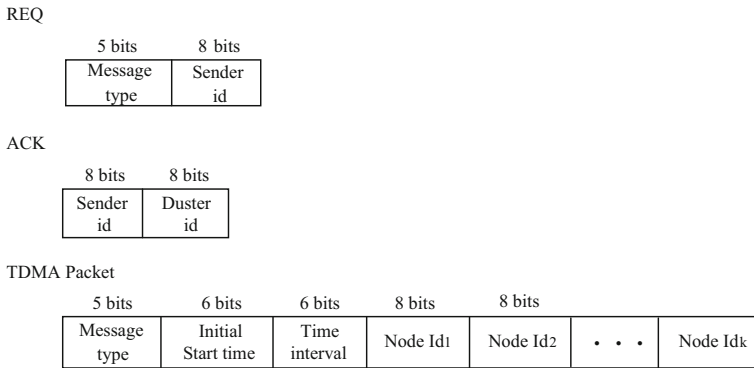


Fig. 3 Data packets

case, aggregate functions such as average function can be effectively while collecting the environmental data. The proposed method employs an average function to aggregate the data in the sensor node.

In most of the methods, the Data Aggregation (DA) is performed in centralized manner such as CH, where each sensor node has to transmit all data to the CH. Since the data are aggregated in the cluster head, the sensor node has unnecessary power consumption in order to transmit the data. In the proposed method, each sensor node employs aggregation function to aggregate the data before it sends to the CH to reduce the amount of traffic. Such that the sensor node collects and aggregates the data other than its TDMA time slot. The aggregated data are transmitted to the CH in its TDMA time slot.

Let τ represents the time schedule for all the sensor nodes in the cluster $\{\tau_1, \tau_2, \dots, \tau_k\}$. τ_i denotes the time schedule for a particular sensor node, such that the time interval for the schedule is $\{t_{st}, t_{fh}\}$, where t_{st} and t_{fh} is the start time and finish time of the schedule.

Let k be the number of sensors connected to the Cluster Head. Sensor node S_i aggregates the data sample set $X \{X_1, X_2, X_3, \dots, X_n\}$ taken in the time period T given as $T \forall \tau | \tau_i$. The Aggregation function is given as,

$$P_{avg} = Avg \left(\sum_{j=1}^n X_j \right) \tag{3}$$

The CH stores the aggregated data and transmits to the mobile sink. In order to differentiate the data of the sensor node, the CH uses separate buffer for each sensor node. The data received in particular time slot is stored in the corresponding buffer. The CH transmits the data in each buffer along with the corresponding node ID to the mobile sink.

4 Energy Analysis

Figure shows the proposed data gathering architecture implemented in a square sensor region. The sensor region is covered by six trace paths, where a single mobile sink is used for two trace paths. Let N is the total number of sensor node present in whole sensor region, where the separated region between the trace paths has ‘ n ’ number of sensor node. The separated sensor region is similar to the rectangular region. The cluster formed has k number of sensors, such that the number of clusters in the separated region is n/k .

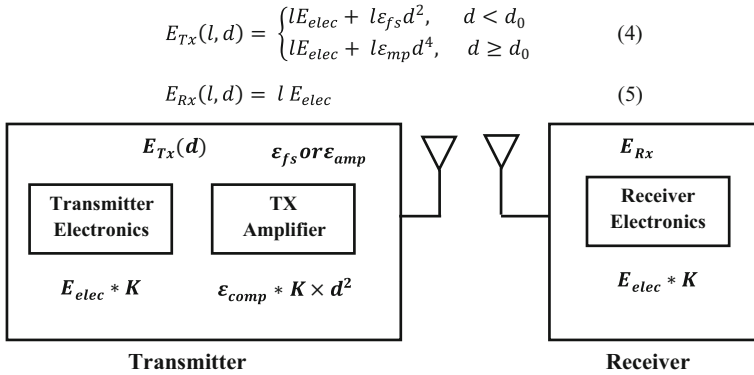


Fig. 4 Radio dissipation model

The sensor nodes equally transmit M bits of data. The energy required to transmit and receive ‘l’ bits of data for distance dⁿ based on the Radio Energy Dissipation model of sensor node as shown in Fig. 4 is E_{Tx}(l, d) and E_{Rx}(l, d) respectively, such that

$$E_{Tx}(l, d) = \begin{cases} lE_{elec} + l\epsilon_{fs}d^2, & d < d_0 \\ lE_{elec} + l\epsilon_{mp}d^4, & d \geq d_0 \end{cases} \quad (4)$$

$$E_{Rx}(l, d) = lE_{elec} \quad (5)$$

The energy required in transmission depends upon the path loss exponent n, where n is 2 for free space and 4 for multipath interference. E_{elec} is the energy consumed by transmitter/receiver electronics upon single bit. ‘d’ is the distance between the transmitter a receiver. The energy consumption of transmitter amplifier is given by ε_{fs}, ε_{mp} based on the path loss exponent. Since the sensor nodes perform aggregation, the ‘m’ bits of data are aggregated into fewer bits of ‘h’. The energy consumption of the model with centralized data aggregation (i.e. Cluster Head) is given by E_{C_DA}. The energy consumption of the proposed method for Data Aggregation performed in the sensor node is given by E_{S_DA}. E_{C_DA} and E_{S_DA} denotes energy used by cluster for both transmitting and receiving the data. The energy gain E_{gain} for a cluster achieved by proposed method is,

$$E_{S_DA} = m(E_{elec} + \epsilon_{fs}d^2) + mE_{elec} \quad (6)$$

$$E_{C_DA} = h(E_{elec} + \epsilon_{fs}d^2) + hE_{elec} \quad (7)$$

$$E_{gain} = E_{S_DA} - E_{C_DA} \quad (8)$$

$$E_{gain} = (2m - h)E_{elec} + (m - h)\epsilon_{fs}d^2 \quad (9)$$

Since, there are (N/k) clusters in the sensor region. The Total energy gain of the proposed method is given by,

$$E_{gain_Tot} = \left(\frac{N}{k}\right) E_{gain} \quad (10)$$

5 Performance Evaluation

In order to evaluate the efficiency, the proposed algorithm is simulated in a 500×500 m square region with 500 nodes. Table gives the parameter value of radio model. The heterogeneous nodes are 10 % of the total nodes, whereas its energy level is three times of the normal nodes. The nodes are uniformly distributed in the sensor region, where the heterogeneous nodes are distributed along the trace path. Simulation parameters are shown in Table 1.

The proposed algorithm is compared with the normal mobile sink based data gathering approach, which uses the same trace path for data collection. The data reception capabilities are analyzed with varying the mobile speed as 100 and 200 cm/s. Comparison of the amount of data packets received for sink mobility speed of 100 cm/s is shown in Fig. 5. Due to the ordered data gathering approach along with data aggregation reduces the number of packets to sense the same region. With sink mobility speed of 100 cm/s, the mobile sink covers a single trace path in 500 s, such that the pair of trace path is covered by 1000 s. Comparison of the amount of data packets received for sink mobility speed of 200 cm/s is shown in Fig. 6. With sink mobility speed of 200 cm/s, the mobile sink covers a single trace path in 250 s, such that the pair of trace path is covered by 500 s. When the speed of the mobile sink is increased, the mobile sink covers the sensor region with increase in number of times, receiving more packets.

Since the sensor nodes aggregates the data before sending to the Cluster Head, reduces the amount of traffic maintain the residual energy of nodes. Further the heterogeneous nodes are only allowed to act as Cluster Head, which helps to retain the total residual energy of the network. This helps the proposed algorithm to achieve network lifetime

Table 1 Parameter of the radio model

Parameter	Value
E_{elec}	50 nJ/bit
ϵ_{fs}	10 pJ/bit/m ²
ϵ_{mp}	0.0013 pJ/bit/m ⁴
E_{DA} (for data aggregation)	5 nJ/bit/signal

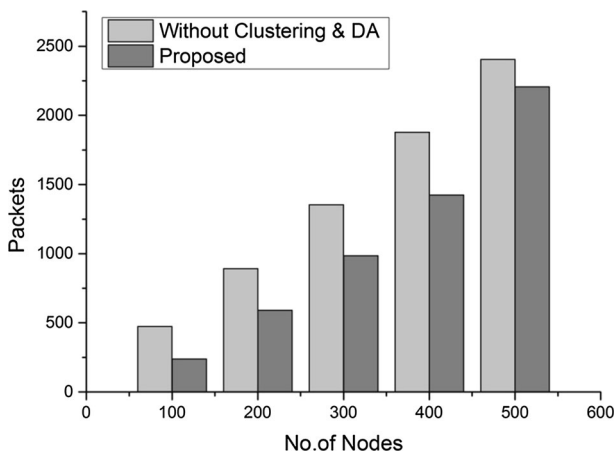


Fig. 5 Number of data packet received with sink speed of 100 cm/s

Fig. 6 Number of data packet received with sink speed of 200 cm/s

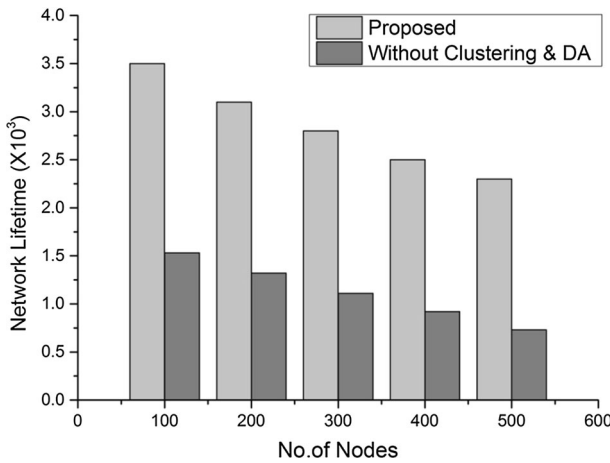
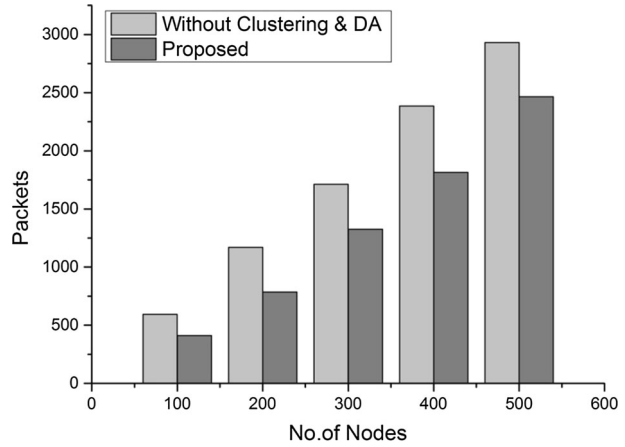


Fig. 7 Comparison of network lifetime by varying the number of nodes

increase compared to other data gathering approach as shown in Fig. 7. The proposed method has more overall residual energy compared to other algorithm as shown in Fig. 8. When the node density (i.e. increase in number of nodes), the residual energy of the overall network is increased (Fig. 8).

Figure 9 shows that the delay performance by varying the number of nodes. Compared to the existing scheme, proposed technique has lesser delay. Figure 10 shows that the throughput analysis. Channel bandwidth is assigned for the proposed scheme is 2 Mbps and the proposed technique achieves 1.4 Mbps throughput by varying the number of nodes. Figure 11 shows that the delivery rate performance. Proposed technique has 40–60 % packets are delivered by varying the number of nodes. Remaining packets are dropped, that packet drop rate analysis is shown in Fig. 12 by varying the number of nodes from 100 to 500 nodes.

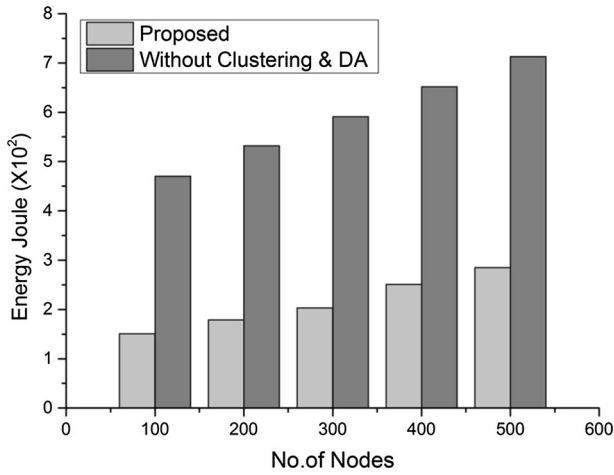


Fig. 8 Comparison of energy usage by varying the number of nodes

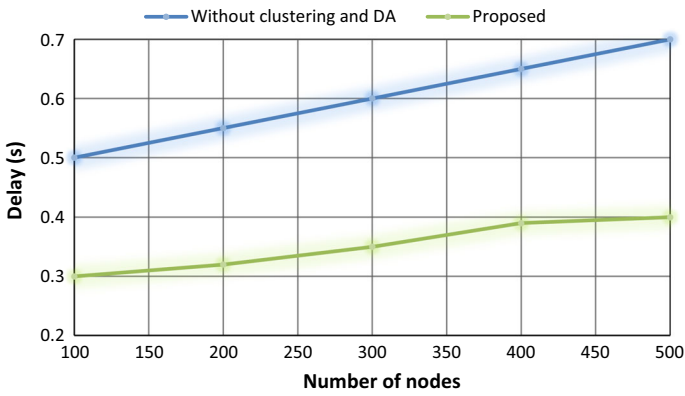


Fig. 9 Comparison of delay by varying the number of nodes



Fig. 10 Comparison of throughput by varying number of nodes

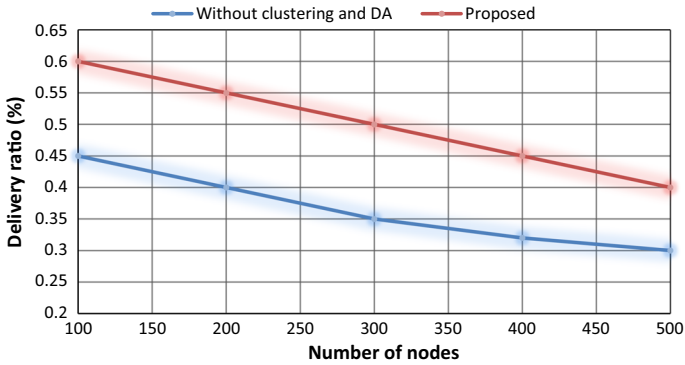


Fig. 11 Comparison of delivery rate by varying number of node

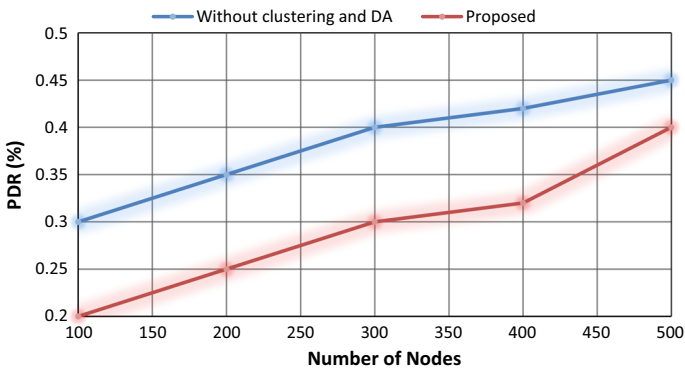


Fig. 12 Comparison of packet drop rate by varying number of nodes

6 Conclusion

In the proposed algorithm, data gathering based on the clustering architecture with TDMA time slot helps to achieve an efficient data gathering approach. Further the data aggregation on the nodes within the cluster reduces the data traffic. The efficiency of the proposed model is evaluated based on number of data packets received, network lifetime and residual energy. The proposed algorithm achieves network lifetime increase with low energy utilization compared to other algorithm. The aim is not to highlight ourself but give a solution for network researchers and motivate to work in a new direction.

References

1. Akyildiz, I. F., Su, W., Sankarasubramaniam, Y., & Cayirci, E. (2002). Wireless sensor networks: A survey. *Computer Networks*, 38, 393–422.
2. Akkaya, K., & Younis, M. (2005). A survey on routing protocols in wireless sensor networks. *Ad Hoc Networks*, 3, 325–349.

3. Heinzelman, W. R., Chandrakasan, A., & Balakrishnan, H. (2000). Energy-efficient communication protocol for wireless sensor networks. In *Proceedings of the 33rd annual Hawaii international conference on system sciences*, Washington, DC, USA (pp 1–10).
4. Heinzelman, W. B., Chandrakasan, A. P., & Balakrishnan, H. (2002). An application-specific protocol architecture for wireless microsensor networks. *IEEE Transactions on Wireless Communications*, *1*, 660–670.
5. Hammoudeh, M., & Newman, R. (2015). Adaptive routing in wireless sensor networks: QoS optimization for enhanced application performance. *Information Fusion*, *22*, 3–15.
6. Wang, Z. M., Melachrinoudis, E., & Basagni, S. (2005). Voronoi diagram-based linear programming modeling of wireless sensor networks with a mobile sink. In *Proceedings of the IIE annual conference and exposition*.
7. Azad, P., & Sharma, V. (2015). Pareto-optimal clustering scheme using data aggregation for wireless sensor networks. *International Journal of Electronics*, *102*(7), 1165–1176.
8. Sert, S. A., Bagci, H., & Yazici, A. (2015). MOFCA: Multi-objective fuzzy clustering algorithm for wireless sensor networks. *Applied Soft Computing*, *30*, 151–165.
9. Mainwaring, A., Polastre, J., Szewczyk, R., Culler, D., & Anderson, J. (2002). Wireless sensor networks for habitat monitoring. In *Proceedings of ACM Int'l workshop wireless sensor networks and applications (WSNA)*.
10. Kaiser, W. J., Pottie, G. J., Srivastava, M., Sukhatme, G. S., Villasenor, J., & Estrin, D. (2003). Networked infomechanical systems (NIMS) for ambient intelligence, Technical report 31, Center for Embedded Networked Sensing, University of California, Los Angeles.
11. Khan, M. I., Gansterer, W. N., & Haring, G. (2012). Static vs. mobile sink: The influence of basic parameters on energy efficiency in wireless sensor networks. *Computer Communications*, *36*, 965–978.
12. Yu, S., Zhang, B., Li, C., & Mouftah, H. T. (2014). Routing protocols for wireless sensor networks with mobile sinks: A survey. *IEEE Communications Magazine*, *52*(7), 150–157.
13. Kansal, A., Somasundara, A., Jea, D. D., Srivastava, M. B. & Estrin, D. (2004). Intelligent fluid infrastructure for energy efficient networks. In *Proceedings of the 2nd ACM/USENIX MobiSys*.
14. Jea, D., Somasundara, A., & Srivastava M. B. (2005). Multiple controlled mobile elements (data mules) for data collection in sensor networks. In *Proceedings of the 1st IEEE/ACM DCOSS*.
15. Intanagonwiwat, C., Govindan, R., & Estrin, D. (2000). Directed diffusion: A scalable and robust communication paradigm for sensor networks. In *MOBICOM* (pp. 56–67).
16. Wichmann, A., & Korkmaz, T. (2015). Smooth path construction and adjustment for multiple mobile sinks in wireless sensor networks. *Computer Communications*. doi:10.1016/j.comcom.2015.06.001.
17. Shi, L., Zhang, B. X., Mouftah, H. T., & Ma, J. (2013). DDRP: An efficient data-driven routing protocol for wireless sensor networks with mobile sinks. *International Journal of Communication Systems*, *26*, 1341–1355.
18. Liu, W., Lu, K. J., Wang, J. P., Xing, G. L., & Huang, L. S. (2012). Performance analysis of wireless sensor networks with mobile sinks. *IEEE Transactions on Vehicular Technology*, *61*, 2777–2788.
19. Lin, C. J. (2006). HCDD: Hierarchical cluster-based data dissemination in wireless sensor networks with mobile sink. In *Proceedings of international conference on wireless communications and mobile computing*, Vancouver, Canada (pp. 1189–1194).
20. Shah, R. C., Roy, S., Jain, S., & Brunette, W. (2003). Data MULEs: Modeling a three-tier architecture for sparse sensor networks. In *Proceedings of IEEE workshop sensor network protocols and applications (SNPA)*.
21. Konstantopoulos, C., Pantziou, G., Gavalas, D., Mpitziopoulos, A., & Mamalis, B. (2012). A rendezvous-based approach enabling energy-efficient sensory data collection with mobile sinks. *Parallel and Distributed Systems, IEEE Transactions on*, *23*(5), 809–817.
22. Small, T., & Haas, Z. (2003). The shared wireless infostation model—A new ad hoc networking paradigm (or where there is a whale, there is a way). In *Proceedings of ACM MobiHoc*.
23. Heidemann, J., Silva, F., Intanagonwiwat, C., Govindan, R., Estrin, D., & Ganesan, D. (2001). Building efficient wireless sensor networks with low-level naming. In *SOSP*. Oct 2001.
24. Intanagonwiwat, C., Estrin, D., Govindan, R., & Heidemann, J. (2001). Impact of network density on data aggregation in wireless sensor networks. Submitted for Publication, ICDCS-22.
25. Intanagonwiwat, C. Govindan, R., & Estrin, D. (2000). Directed diffusion: A scalable and robust communication paradigm for sensor networks. In *MobiCOM*, Boston, MA.



A. Muthu Krishnan is working as a faculty in Electronics and communication engineering in Regional centre of Anna University Madurai. He completed his graduation from Anna University and his Master degree in VLSI Design from Anna University of Technology, Coimbatore. He had 5 years of Teaching experience in Electronics and communication Engineering. Presently he is doing his Ph.D. in Anna University, Chennai. Currently he is published 25 research papers in reputed International, national journal and conferences.



P. Ganesh Kumar obtained his Bachelor of Engineering from Madurai Kamaraj University and received his Master of Engineering in Applied Electronics from Bharathiyar University in 2000. Also he obtained his Doctoral (Ph.D.) degree in Information and Communication Engineering from Anna University Chennai in 2009. He has served at various reputable institutions and industries in India and China. Currently he is working as a senior Technical consultant in Linyi Top Network Co Ltd, Linyi city, Shangdong Province, China from June 2014. Also working as a Professor in M.E Networking in K.L.N College of Engineering from June 2013. Currently he is holding three patents and published 95 research papers in reputed International, national journal and conferences. He is a grant recipient of Tamil Nadu State council for Science and Technology (TNSCST) and Motorola Scholar award, Tata Consultancy Services (TCS) and All India Council for Technical Education (AICTE). He has been invited as Chairpersons and delivered special lectures in many international and national conference and workshops. His current areas of interest are Internet of Things (IoT) and Wireless networks. He is a reviewer for many international, national journals and conferences. He is a recognized supervisor for M.S (by research) and Ph.D. by Anna University Chennai and now 11 students are doing Ph.D. under him.